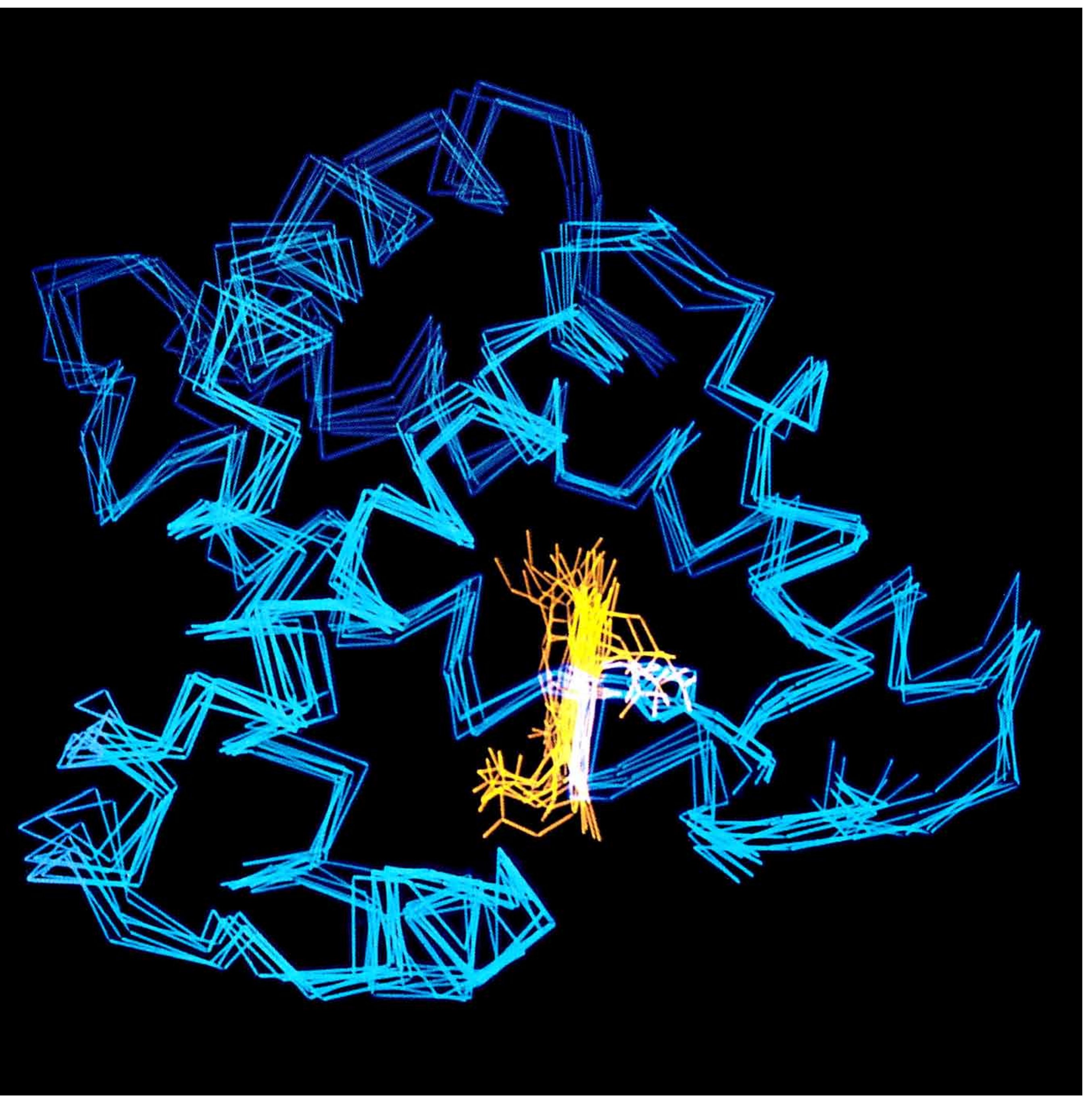
INVESTIGACION Y CIENCIA

Edición en español de **SCIENTIFIC AMERICAN**



DINAMICA DE LAS PROTEINAS

Junio 1986

450 PTAS.
(IVA INCLUIDO)

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 117
Junio 1986

INVESTIGACION Y
CIENCIA
Edición en español de SCIENTIFIC AMERICAN

SCIENTIFIC AMERICAN

COMITE DE REDACCION Jonathan Piel (Presidente y director), Timothy Appenzeller, John M. Benditt, Peter G. Brown, David L. Cooke, Jr., Ari W. Epstein, Michael Feirtag, Gregory R. Greenwell, Robert Kunzig, Philip Morrison, James T. Rogers, Armand Schwab, Jr., y Ricki L. Rusting

DIRECCION ARTISTICA Samuel L. Howard
PRODUCCION Richard Sasso
DIRECTOR GENERAL George S. Conn

INVESTIGACION Y CIENCIA

DIRECTOR Francisco Gracia Guillén
REDACCION José María Valderas Gallardo (Redactor Jefe)
Carlos Oppenheimer
José María Farré Josa
PRODUCCION César Redondo Zayas
PROMOCION EXTERIOR Pedro Clotas Cierco
EDITA Prensa Científica, S.A.
Calabria, 235-239
08029 Barcelona (ESPAÑA)

Colaboradores de este número:

Asesoramiento y traducción:

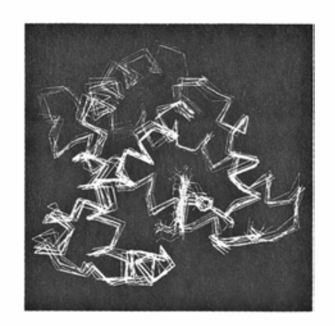
Emilio Elizalde Rius: *Aplicaciones del efecto Hall cuántico*; Manuel Crespo Veigas: *Ritmos internos en la emigración de las aves*; Jorge Peña Gutiérrez: *Nuevas estrategias para la energía nuclear*; José M. Vidal Llenas: *Giroscopios ópticos*; Natividad Sánchez Sáinz-Trápaga: *Las "tensiones" del nacimiento*; Andrés Lewin Richter: *Interpretación de la trompeta barroca*; Luis Bou: *Juegos de ordenador*; J. Vilardell: *Taller y laboratorio*.

Ciencia y sociedad:

Luis de Medinaceli

Libros:

Tito Antonio Varela, Julio A. Alonso Martín, Antoni Roca, Antonio Fernández-Rañada y Mariano Artigas



PORTADA

La ilustración de portada es una composición de siete imágenes distintas, que abarcan un intervalo de 30 billonésimas de segundo de una simulación por ordenador de los movimientos internos de la proteína denominada mioglobina (véase "Dinámica de las proteínas", por Martin Karplus y J. Andrew McCammon, en este mismo número). Si careciera de movimiento interno, la mioglobina no podría desempeñar la función que le corresponde, a saber, el almacenamiento de oxígeno en el tejido muscular animal: los movimientos permiten al oxígeno abrirse camino por entre el denso empaquetamiento atómico de la molécula hasta su región de unión, una compleja molécula orgánica plana, denominada grupo hemo, que se muestra en naranja. La perspectiva elegida lo fue para que se apreciara la mioglobina desde un ángulo bajo el cual aparece relativamente poco empaquetada. El ordenador permite rotar la molécula a voluntad.

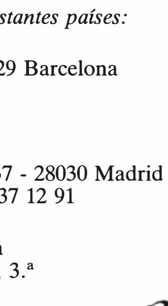
Suscripciones:
Prensa Científica, S. A.
Calabria, 235-239
08029 Barcelona (España)
Teléfono 322 05 51 ext. 33-37

Condiciones de suscripción:
España:
Un año (12 números):
4400 pesetas (IVA incluido)
Extranjero:
Un año (12 números): 33 U.S. $
Ejemplar atrasado ordinario:
450 pesetas (IVA incluido)
Ejemplar atrasado extraordinario:
575 pesetas (IVA incluido)

Distribución para España:
Distribuciones de Enlace, S. A.
Bruch, 49 - 08009 Barcelona

Distribución para los restantes países:
Editorial Labor, S. A.
Calabria, 235-239 - 08029 Barcelona

Publicidad:
Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137 - 28030 Madrid
Teléfonos 430 84 81 - 437 12 91
Cataluña:
Paulino Muñoz Victoria
Muntaner, 467-469, 3.°, 3.ª
08021 Barcelona
Teléfono 211 92 45



Controlado
por O.J.D.

PROCEDENCIA DE LAS ILUSTRACIONES

Fotografía de la portada de James Kilkelly

Página	Fuente	Página	Fuente
8	R. Díaz de la Iglesia		Nacional de la Salud Mental
10-14	R. Díaz de la Iglesia y		de los EE.UU.
	Joan Cotoner	49-55	Tom Prentiss
11	R. Díaz de la Iglesia (*arriba*)	62-69	Ilil Arbel
15	R. Díaz de la Iglesia	72-73	Jon Brenneis
16	R. Díaz de la Iglesia y	74-76	Ian Worpole
	Joan Cotoner	77	Jon Brenneis
19	James Kilkelly	81	Charles R. Belinky, Photo
20-21	Ilil Arbel		Resarchers, Inc.
22-25	James Kilkelly	82-87	Patricia J. Wynne
26	Carl Bränden, Universidad de	88	Thomas Bergman
	Uppsala (*arriba*); Sally Black	91	Don Smithers
	(*abajo*)	92-96	George Retseck
27	Axel Brüngen, Universidad	99	Thomas Banchoff, David
	de Harvard (*arriba*); Edward		Laidlaw y David Margolis,
	Bell (*abajo*)		Universidad de Brown
31-37	Hank Iken, Walken Graphics	100	Ilil Arbel (*arriba*), Ian
38	Horst L. Störmer, AT&T Bell		Worpole (*abajo*)
	Laboratories	101-102	Thomas Banchoff, Hüseyin
42	A. L. van Beck en *Plastic and*		Koçak, David Laidlaw y
	Reconstructive Surgery, pág.		David Margolis, Universidad
	510, 1979		de Brown
43-45	Luis de Medinaceli, Saint	103	Ilil Arbel
	Elizabeths Hospital, Instituto	106-110	Michael Goodman

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición Tecfa, S.A.
Pedro IV, 160 - 08005 Barcelona
Fotocromos reproducidos por GINSA, S.A.
Gráfica Elzeviriana, S.A.
Nápoles, 249 - Tel. 207 40 11
08013 Barcelona
Printed in Spain - Impreso en España

Los autores

RAIMUNDO DIAZ DE LA IGLESIA ("Comunicaciones ópticas por cables submarinos") trabaja en la Compañía Telefónica Nacional de España (CTNE). Durante el período 1977-1980, fue profesor de teoría de la información y codificación de la Facultad de Informática de la Universidad Politécnica de Madrid. Desde 1978, fecha en que ingresa en la CTNE, centra su principal actividad en aspectos relacionados con la normalización de vías de transmisión de gran capacidad y, particularmente desde finales de 1979, en vías de transmisión por fibra óptica; sobre esta materia ha publicado artículos y presentado ponencias en congresos nacionales e internacionales. Es miembro del *IEEE*.

MARTIN KARPLUS y J. ANDREW McCAMMON ("Dinámica de las proteínas") son, respectivamente, catedrático Theodore William Richards de química de la Universidad de Harvard y catedrático M. D. Anderson de química de la de Houston. Nacido en Austria, Karplus se licenció en química y física por Harvard, en 1951, estudiando luego quimiofísica con Linus Pauling en el Instituto de Tecnología de California, por el que se doctoró en 1954. Pasó dos años en la Universidad de Oxford trabajando con Charles A. Coulson, incorporándose luego al claustro docente de la Universidad de Illinois en Urbana-Champaign. Karplus se trasladó a la Universidad de Columbia en 1960, y pasó a ocupar su plaza actual en Harvard en 1965. McCammon se formó en química y física en el Pomona College y en Harvard, institución que le otorgó el doctorado en 1976, con una tesis dirigida por Karplus y John M. Deutch.

BERTRAND I. HALPERIN ("Aplicaciones del efecto Hall cuántico") es catedrático de física de la Universidad de Harvard. En ese centro superior inició su formación, terminándola en la Universidad de California en Berkeley, por la que se doctoró en 1965. El año siguiente lo pasó en la Universidad de París becado por la Fundación Nacional de Ciencias de los Estados Unidos. Posteriormente trabajó en los Laboratorios Bell de Murray Hill, New Jersey. Desde 1973 hasta 1980 fue redactor de la revista *Reviews of Modern Physics*. En 1982 Halperin recibió el premio Oliver E. Buckley de Física de la Materia Condensada que otorga la Sociedad Americana de Física.

EBERHARD GWINNER ("Ritmos internos en la emigración de las aves") dirige el Vogelwarte Radolfzell, estación ornitológica del Instituto Max Planck de Fisiología del Comportamiento, y es profesor adjunto de fisiología de la Universidad de Munich. La de Tübingen le concedió el doctorado, en 1964, por su tesis sobre el comportamiento social del cuervo. Se centró luego en la investigación de diversos aspectos de los ritmos biológicos del comportamiento de las aves, en el instituto dirigido por Jürgen Aschoff en Erling-Andechs, en la Universidad de Washington y en la de Stanford. Los trabajos de Gwinner se realizan hoy principalmente en el laboratorio, pero no por ello deja de considerarse un biólogo de campo.

RICHARD K. LESTER ("Nuevas estrategias para la energía nuclear") es profesor de ingeniería nuclear del Instituto de Tecnología de Massachusetts (MIT). Nacido en Inglaterra, se licenció por el Imperial College de Ciencia y Técnica de Londres. Becado por el Kennedy Memorial Trust, en 1974 marchó a los Estados Unidos para seguir estudios de ingeniería nuclear en el MIT. En 1977, durante año y medio de excedencia del MIT, fue becario investigador visitante en la división de relaciones internacionales de la Fundación Rockefeller. En 1979 se doctoró por el MIT, incorporándose a su claustro en calidad de profesor ayudante; en 1982 ocupó plaza de profesor asociado.

DANA Z. ANDERSON ("Giroscopios ópticos") da clases de física en la Universidad de Colorado en Boulder. Pertenece al Instituto de Astrofísica de la misma. Titulado en ingeniería eléctrica por la Universidad de Cornell (1975), se doctoró en física por la de Arizona, en 1981. En calidad de becario pstodoctoral del Instituto de Tecnología de California colaboró en el desarrollo de un interferómetro detector de las ondas de gravedad. La preocupación de Anderson por el distanciamiento entre ciencia y humanidades le llevó a colaborar con Eric White, del departamento de filología inglesa de la Universidad de Colorado, en la organización de un curso titulado "Contemplando la naturaleza: las implicaciones humanísticas de la física moderna".

HUGO LAGERCRANTZ y THEODORE A. SLOTKIN ("Las tensiones del nacimiento") comparten un mismo interés por el desarrollo del recién nacido. Lagercrantz es profesor asociado de fisiología y perinatología del Instituto Carolino de Estocolmo, por el que se doctoró en fisiología bajo la tutoría del desaparecido Ulf von Euler. A partir de 1976 se ha ocupado fundamentalmente del estudio de la liberación y función de las catecolaminas (que actúan como hormonas y como agentes neurotransmisores) en bebés. Slotkin es profesor de farmacología y psiquiatría de la Facultad de Medicina de la Universidad de Duke. Licenciado en química por el Brooklyn College en 1967, se doctoró en farmacología y toxicología (1970) por la Universidad de Rochester. Durante los últimos diez años ha investigado intensamente la bioquímica, fisiología y farmacología del desarrollo del sistema nervioso, así como la regulación de la función cardiovascular en el recién nacido. En 1976 Slotkin recibió el premio de la Asociación Cardiaca de Carolina del Norte al investigador joven más destacado y, en 1982, el John Jacob Abel, que otorga la Asociación Americana de Farmacología y Terapéutica Experimental.

DON SMITHERS, KLAUS WOGRAM y JOHN BOWSHER ("Interpretación de la trompeta barroca") aportan al estudio de la trompeta sus propios conocimientos, muy distintos. Smithers enseña historia de la música y de los instrumentos musicales, y es miembro de la Real Asociación británica de Música. Estudió en la Universidad de Hofstra, licenciándose en 1957. En 1967 se doctoró en historia de la música por la Universidad de Oxford. Fue profesor de la Universidad de Syracuse y del Real Conservatorio de Música de Holanda. Smithers es el intérprete de una fanfarria para trompeta barroca que se emite en la sintonía de un noticiario de la CBS. Wogram trabaja en la división de física acústica de la Oficina alemana de Pesos y Medidas. Se licenció en ingeniería eléctrica por la Politécnica de Braunschweig, en 1967, doctorándose en acústica física por esa misma institución en 1972. Bowsher enseña acústica en la Universidad de Surrey y es vicepresidente del instituto de acústica. Se licenció y doctoró en la facultad de Ciencia y Técnica de Londres en 1957. Antes de incorporarse a su cargo actual investigó en la sección de acústica del Laboratorio Nacional de Física de Middlesex.

Hace...

José M.ª López Piñero

...cuatrocientos años

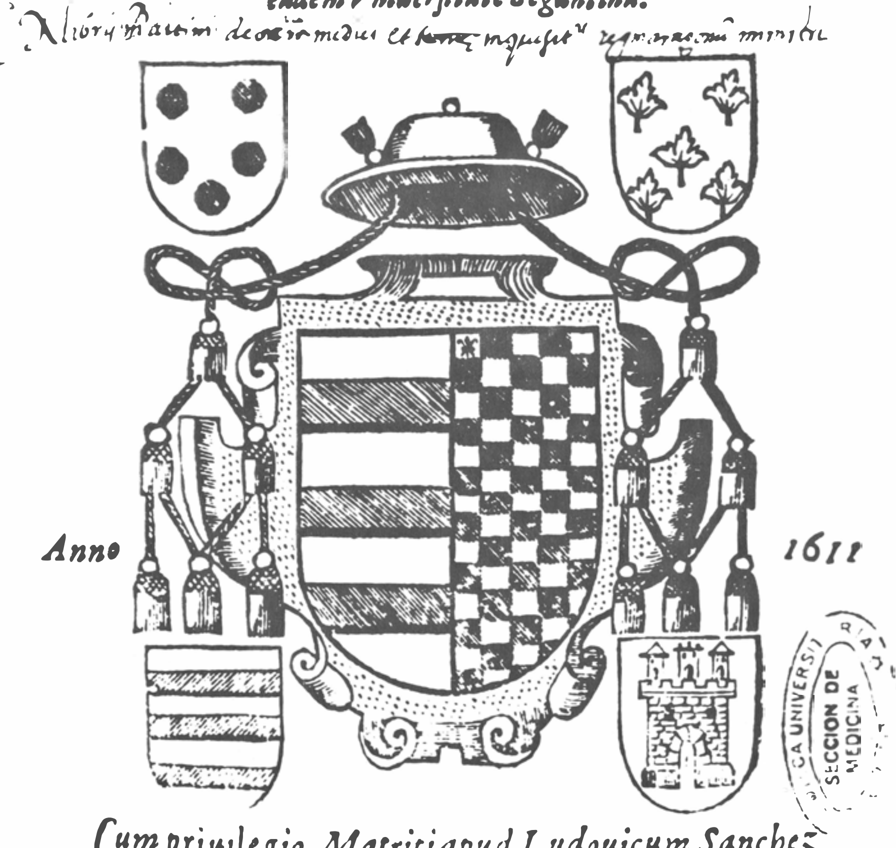
Francisco Pérez Cascales participó en las oposiciones a una de las cátedras médicas de la Universidad de Alcalá, resultando envuelto en un escándalo por haberse visto implicado en algunas gestiones irregulares. En consecuencia, tuvo que abandonar Alcalá y renunciar de momento a su carrera universitaria. Nacido en Guadalajara, se había formado en el período de mayor esplendor de la medicina universitaria complutense, obteniendo el título de licenciado en 1579 y el de doctor el año siguiente.

Tras el incidente anotado, Pérez Cascales residió durante quince años en Yepes, donde fue nombrado médico del duque de Maqueda tras haber tratado con éxito la enfermedad de uno de sus hijos.

En 1601 se trasladó a Sigüenza, donde fue médico de su cabildo diocesano y titular de la cátedra de prima de medicina de su Universidad que, como es sabido, era una de las llamadas "menores". Diez años más tarde publicó su única obra generalmente considerada como una de las más importantes aportaciones españolas a la literatura pediátrica de la época: *Liber de affectionibus puerorum, una cum tractatu de morbo illo vulgariter Garrotillo appellato, cum duabus Quaestionibus*. La dedicó a Antonio Venegas Figueroa, influyente personaje eclesiástico, entonces obispo electo de Sigüenza. Es un volumen en octavo de casi trescientas páginas, que comprende cincuenta capítulos consagrados a las enfermedades de los niños y cuatro apéndices de carácter monográfico. A diferencia de otros tratados pediátricos de este período, no se ocupa de cuestiones de puericultura y tampoco ofrece capítulos introductorios sobre patología general infantil. La mayor parte de las afecciones que estudia son las que aparecen en los textos de la tradición "pediátrica" de origen clásico. Las expone en la habitual ordenación de la cabeza a los pies: aftas bucales, epilepsia, "estupor" o abolición incompleta de la sensibilidad y la motilidad, pústulas del cuero cabelludo, parálisis, rabia o "espasmo cínico", hidrocefalia, tos, estornudos molestos, sanies o pus en los oídos, sanies o pus en los ojos, estrabismo, conjuntivitis, "nubécula" y cicatrices corneales, hipo, vómitos, tumefacción del ombligo, supresión de la orina, cálculos urinarios, intertrigo, sabañones, viruela y sarampión, lombrices y gusanos.

A grandes rasgos, el esquema nosográfico de Pérez Cascales coincide con el del médico renacentista italiano Paolo Bagellardo y el del gran clínico islámico medieval Rhazes. En cambio, el contenido de sus capítulos es radicalmente distinto. Concede gran atención al estudio de la "naturaleza", la localización, las causas y la patogenia de cada afección, de acuerdo con las teorías de la patología del galenismo tradicional, sin concesión alguna al paracelsismo o a cualquier otra corriente renovadora. En este punto, el profesor seguntino puede encuadrarse en el llamado "galenismo contrarreformista" que suele personificarse en la obra de Luis Mercado, aunque entre ambos autores existen algunas diferencias que interesa anotar.

La primera de ellas es que Pérez Cascales carece de la ambición sistemati-



LIBER
DE AFFECTIONIBVS
PVERORVM, VNA CVM
TRACTATV DE MORBO ILLO VVLGARITER
Garrotillo appellato, cum duabus Quæstionibus. Altera, De gerentibus vtero rem appetentibus denegatam.
Altera verò de Fascinatione.

PER DOCTOREM FRANCISCVM PEREZ
Cascales de Guadalajara Decani, & Capituli sanctæ Ecclesiæ Seguntinæ Medicum, & Primarium Medicinæ professorem in eadem Vniuersitate Seguntina.

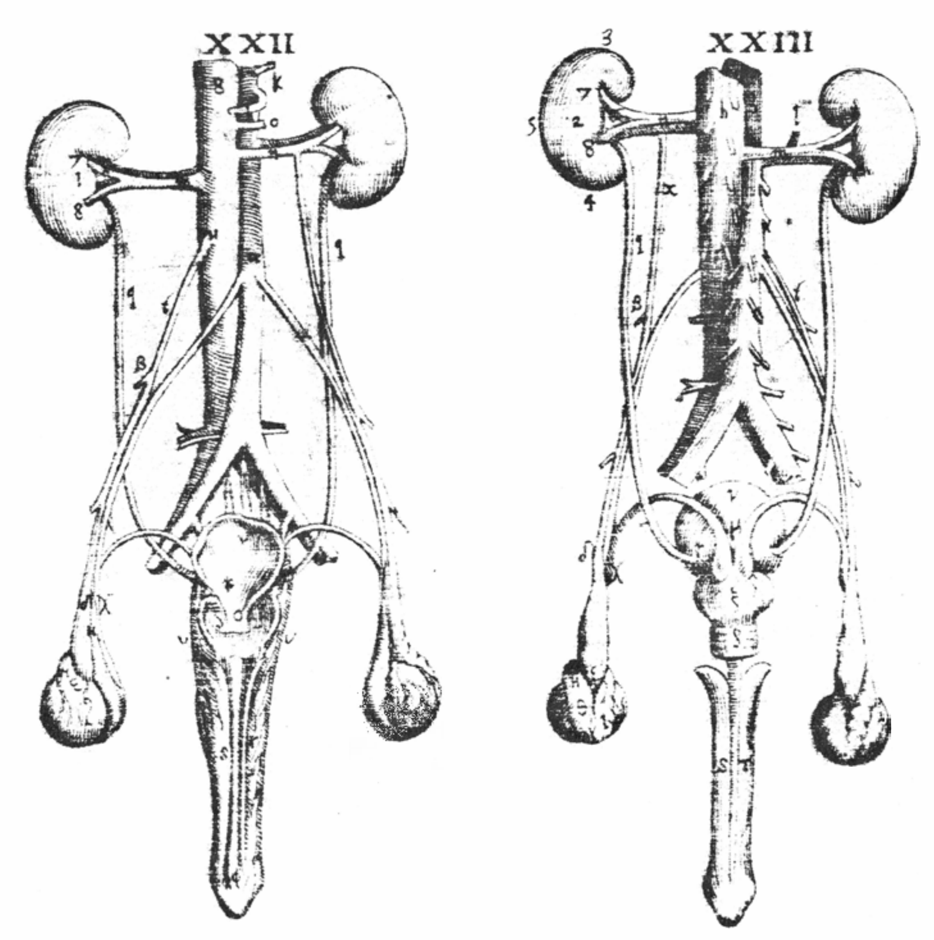
Anno 1611

Cum priuilegio. Matriti apud Ludouicum Sanchez Typographum Regium.

1. Portada del tratado de Pérez Cascales.

zadora que informa toda la obra de Mercado y, por lo tanto, de su intención casi puramente doctrinal. Por el contrario, su libro responde a un equilibrio entre teoría y práctica que cuadra a su condición de profesor universitario que al mismo tiempo era un experimentado clínico. En el prólogo, aduce a modo de justificación: "He estado dedicado a la profesión médica durante treinta y cuatro años y, con el auxilio de Dios omnipotente, he asistido y devuelto a su primitiva salud un número casi incontable de lactantes y de niños que padecían gravísimas y diversas enfermedades." Contaba asimismo con una excelente formación libresca, que se refleja en las abundantes y oportunas citas que incluye en su obra. Ofrece incluso una lista previa de sesenta y dos autores citados en la que, aparte de las "autoridades" antiguas y medievales, figura un elevado número de médicos del siglo XVI, entre ellos las cabezas de las distintas tendencias del galenismo en España, Italia y Francia. Por su formación en la Universidad de Alcalá, Pérez Cascales estaba especialmente influido por el llamado "galenismo hipocratista", que habían encabezado en las aulas complutenses Francisco Valles y Cristóbal de Vega y que, a principios del siglo XVII, mantenía en ellas principalmente el magisterio de Pedro García Carrero. Dicha tendencia se caracterizaba por asumir el galenismo, pero concediendo gran importancia a la observación clínica –de la que Hipócrates se consideraba modelo– y a los datos anatómicos procedentes de la disección de cadáveres humanos, que se utilizaban para fundamentar o rectificar las teorías tradicionales sobre la patogenia y la localización de las enfermedades. Ambos elementos están presentes en la obra de Pérez Cascales, que recurre con frecuencia a su propia casuística, exponiendo como base de sus argumentos excelentes historias clínicas, y tiene en cuenta los datos anatómicos en la mayoría de sus exposiciones patogénicas.

Los cuatro apéndices del libro de Pérez Cascales son un "Tractatus de morbo illo vulgariter Garrotillo appellato", dos "quaestiones" sobre el ayuno y las sangrías como posibles causas de aborto y una tercera "Quaestio de fascinatione". El primero figura entre los "clásicos españoles del garrotillo", es decir, la contribución hispánica a la nosografía de la angina diftérica sofocante. La última es un interesante ejemplo de la lucha que, en esta época, se dirigió desde la cultura académica contra las "supersticiones y hechicerías", en este caso, contra las tocantes a la fascinación o mal de ojo como presunta causa de enfermedades.



2. Para estudiar la "supresión de la orina sin que llegue a la vejiga", Pérez Cascales, sin abandonar las teorías galénicas tradicionales, se apoya en la observación clínica y en los datos anatómicos. Estos esquemas del aparato génitourinario masculino proceden del tratado de anatomía (1556) de Juan Valverde de Amusco.

Entre las contribuciones del libro de Pérez Cascales se han destacado las relativas a las dos afecciones que estudia más ampliamente: la epilepsia y la supresión de la orina.

Su exposición sobre la epilepsia comienza con una erudita introducción sobre los distintos nombres de la enfermedad, adelantando ya su condición de enfermedad cerebral. Expone a continuación la "esencia" de la misma conforme a las teorías humoralistas del galenismo. Defiende en concreto que consiste en la coincidencia de una materia obstructiva y una causa irritante que ocasionan convulsión de todas las partes del cuerpo y que aparece a intervalos acompañada de la pérdida de los sentidos externos e internos. Los sentidos internos eran, para un médico renacentista, la imaginativa, la razón y la memoria; según la teoría aceptada por Pérez Cascales, resultaban anulados cuando un humor espeso obstruía los ventrículos del cerebro, impidiendo el paso de los "espíritus animales". Estos últimos eran sustancias muy tenues encargadas de transmitir a todo el cuerpo, a través de los nervios, la acción de la "facultad animal" responsable de las sensaciones y los movimientos; la palabra "espíritu" no tiene en este término (ni en el de "espíritus naturales" que ya hemos usado) el sentido de "alma inmortal", sino el de "vapor sutilísimo como el que (se supone) exhalan los licores", llamados por ello "espirituosos". Para explicar las convulsiones, Pérez Cascales defiende una "mordacidad" o "acrimonia" irritante, o sea, una alteración del propio humor obstructivo que lo convierte en acre y corrosivo.

Pérez Cascales distingue tres tipos de epilepsia, el primero de los cuales considera que es una afección primaria del cerebro, mientras que los otros dos serían enfermedades por "consenso", bien con el estómago, bien con otros órganos. El concepto de "consenso", de gran importancia en la patología galénica, implicaba que el padecimiento de una determinada parte del cuerpo podía pasar a otros a distancia, a través de diferentes vías. Pérez Cascales se esforzó principalmente en exponer los síntomas que permitían un diagnóstico diferencial entre los tres tipos.

En su estudio sobre la epilepsia, Pérez Cascales cita fundamentalmente obras hipocráticas y galénicas. Del *Corpus Hippocraticum*, tres de los libros que la moderna crítica incluye en el grupo de los genuinamente coicos. Uno de ellos es, naturalmente, *Sobre la enfermedad sagrada*, la monografía hipocrática sobre la epilepsia, generalmente considerada como un hito en la superación del pensamiento mágicorreligioso en medicina; los otros dos son el célebre estudio ambientalista *Sobre los aires, las aguas y los lugares* y las colecciones de historias clínicas tituladas *Epidemias*. De Galeno, las obras citadas son tres de sus más importantes escritos patológicos (*Sobre las diferencias de las enfermedades, Sobre los lugares afectos, Sobre las causas de las enfermedades*) y los comentarios a las *Epidemias* hipocráticas. Las otras referencias corresponden a los *Problemas*, texto misceláneo seudoaristotélico, al autor latino bajomedieval Celio Aureliano, a los bizantinos Pablo de Egina, Aecio y Alejandro de Tralles y al *Canon* del árabe Avicena.

El estudio de Pérez Cascales sobre "la supresión de la orina sin que llegue a la vejiga" gira en torno a su experiencia clínica personal. Entre las cuestiones patológicas que plantea, destaca la rivalidad a la posibilidad de que se suprima totalmente la orina por obstrucción de un solo riñón o de un solo uréter. Expone y refuta la teoría que lo explicaba mediante la alteración por "consenso" del riñón o uréter del otro lado, así como la de su maestro Francisco Valles, que suponía, entre otras cosas, la participación de los músculos abdominales en la expulsión de la orina. Frente a ellas, defiende un mecanismo basado en la obstrucción del cuello vesical por descenso de arenillas o de humores espesos. Al exponer el tratamiento, vuelve a describir los casos clínicos con mayor detalle, anotando junto a cada síntoma las interpretaciones que le habían servido para sentar las indicaciones terapéuticas. En una primera fase, consistieron éstas en un enema y una sangría como medidas generales y en intentar una acción "emoliente, separativa y digestiva", aplicando localmente una serie de hierbas que enumera y administrando "jarabe de las dos raíces" (de hinojo y apio), que era de efectos diuréticos. En una fase posterior, recurrió a la ingestión de cassia fistula y trementina de abeto, es decir, a un purgante y a un antiséptico urinario.

En este caso, las citas son mucho más escasas. Se limitan a Hipócrates, el romano Celso, el bizantino Pablo de Egina y dos de las más importantes obras médicas españolas del siglo XVI: los *Commentaria* (1577) de Francisco Valles a las *Epidemias* hipocráticas y los *Libri quatuor de Morborum internorum curatione* (1594), de Luis Mercado.

Como ejemplo de la obra de Cascales, he aquí un fragmento de su estudio en torno a la supresión de la orina: "¿Puede la obstrucción de un solo uréter suprimir y cohibir totalmente la orina? Los que piensan que sí afirman que, en los casos que acabamos de exponer, resulta evidente que, al obstruirse y eliminarse uno de los uréteres, la orina fluyó por el otro; y ello lo estiman conforme a la razón pues, del mismo modo que Dios, nuestro creador óptimo y máximo, ha dotado a los animales de dos ojos, de dos fosas nasales y de dos orejas y oídos, para que si les falta un ojo puedan discernir y ver con el otro, o si les falla un oído puedan oír con el otro, el Creador nos ha dotado de dos riñones y de dos vías ureterales, con el fin de que cuando un riñón no funcione o lo haga de manera deficiente le supla el otro y cuando uno

de los uréteres se obstruya no cese por el otro el descenso de la orina desde el riñón a la vejiga. No obstante, hay autores doctísimos que tienen una opinión diferente. Un médico muy sabio y prudente con el que he discutido a menudo acerca de este tema me aseguró que de ninguna manera puede la orina cohibirse y suprimirse sin que se obstruyan ambos riñones, o ambas venas emulgentes o ambos uréteres, ya que si uno de ellos se obstruye por sí, el otro lo hace por consenso. Yo le insistí para que me explicase en qué consiste el consenso que tienen entre sí dichas partes, en virtud del cual, al obstruirse una, la otra puede afectarse, de forma que sufre la obstrucción como la primera, que se ha alterado por sí misma. Respondió que la orina que no puede llegar a la vejiga por el uréter obstruido, regurgitando al otro riñón o al otro uréter, ocluye este último y, así, ambas vías obstruidas impiden el descenso de la orina a la vejiga, produciéndose una supresión completa. Sin embargo, esta respuesta no me convenció ni a mí ni a otros muchos médicos que estaban presentes, ya que no resulta creíble que la orina que regurgita desde la parte obstruida pueda ocluir la otra, puesto que al fluir de modo continuo, a no ser que se impida, no puede de ninguna manera ocluir la vía que la naturaleza ha dispuesto para su propia evacuación...

"Mi maestro el doctísimo Valles, en sus comentarios a las *Epidemias* hipocráticas, parece resolver esta cuestión afirmando que raramente se forma un cálculo en un riñón sin que el otro se obstruya por humores crasos o arenillas, ya que las venas emulgentes tienen su origen común en la vena cava y precisamente por ello la orina se suprime o se expulsa con dificultad al quedar obstruidos ambos riñones. Hasta aquí Valles. Sin embargo, yo planteo a este sabio varón: ¿qué sucede si un uréter queda anulado, porque tiene un cálculo encajado, y la orina se suprime, como sucedió en el enfermo cuya historia clínica hemos expuesto antes? A ello responde Valles que, a causa del dolor, falta el esfuerzo de los músculos del abdomen, que es imprescindible para expulsar la orina... Lo cual no resulta en absoluto convincente por lo siguiente: el enfermo del que antes hemos hablado no padeció un dolor vehementísimo y agudísimo tan pronto como se le obstruyó un uréter por el cálculo; la supresión de la orina coincidió con el dolor y la oclusión de dicho uréter, apareciendo antes de que en la vejiga notara dolor y sensación de peso; en consecuencia, la orina puede suprimirse aunque se conserve el esfuerzo de los músculos del abdomen cuando, al principio, no hay dolor o es pequeño...

"Con el permiso de tantos doctores diré brevísimamente lo que opino acerca de la supresión de la orina que padecieron los enfermos que antes he citado y que, como dije, conseguí curar en la ciudad de Yepes. Pienso que cuando un riñón se obstruye, permaneciendo el otro ileso, y que cuando uno de los uréteres se ocluye, sin que exista lesión en el otro, puede muy bien producirse una supresión de la orina que llegue a ser total. La interpretación que doy a tal fenómeno consiste en que, al bajar el cálculo o los humores crasos que ocasionaban la obstrucción desde los riñones a los uréteres, bajan muchas arenillas o humores crasos que, al llegar al cuello de la vejiga, lo obstruyen, lo cual produce la supresión de la orina. Una vez obstruido el cuello, se impide totalmente la entrada de la orina en la vejiga, presentándose la supresión completa que afectó a los enfermos citados, que tenían un uréter obstruido, pero el otro incólume..."



Comunicaciones ópticas por cables submarinos

Los cables de fibra óptica transportan fotones y sustituyen con ventaja a los cables metálicos. Con ellos se entra en una nueva etapa en la historia del desarrollo de las comunicaciones

Raimundo Díaz de la Iglesia

En los inicios de la era de los satélites de telecomunicación parecía una antigualla hablar de los cables submarinos. Esa opinión estuvo muy generalizada en la década de los sesenta coincidiendo con la euforia de las primeras puestas en órbita. Pero el rendimiento aportado por ambos sistemas y las innovaciones registradas sobre todo en los cables de comunicación submarina permiten asegurar para el futuro la tendencia actual: un equitativo reparto en el tráfico. ¿Cómo explicarlo?

La historia del cable submarino se jalona en tres etapas bien diferenciadas. La primera arranca desde 1851, fecha en que se instaló el primer cable submarino del mundo, que unía Francia con Gran Bretaña. Llega hasta finales de la segunda guerra mundial. Los cables se caracterizaban por su corta vida, inferior a diez años, y por su empleo, casi exclusivo, dedicado a transmitir señales telegráficas.

En 1946 se tendió entre Alemania y Gran Bretaña el primer cable de la que podríamos llamar segunda época. Aportaba como novedad la incorporación de equipos amplificadores intermedios que le facultaban para transmitir canales telefónicos. (El primer cable transatlántico de similares características se instaló en 1956 para unir Gran Bretaña con Estados Unidos.)

La aparición del polietileno, idóneo como dieléctrico y aislante del agua, y el propio desarrollo de la electrónica en los años inmediatos a la terminación de la segunda guerra mundial permitieron prolongar la vida de los cables e intercalar repetidores electrónicos depositados en el suelo marino. Los primeros repetidores eran de válvulas electrónicas, aunque pronto se sustituyeron por equipos transistorizados. Paralelamente a la transistorización, se incluyó, en un mismo repetidor, circuitos que amplificaban las señales enviadas en los dos sentidos de transmisión y, así, bastaba un solo par metálico coaxial para canalizar el tráfico bidireccional. Llámase par metálico coaxial a un conjunto formado por dos conductores metálicos, el primero de los cuales es un cilindro hueco dentro del cual se aloja, alineado axialmente, el otro conductor; se emplean en telecomunicación para enviar altos volúmenes de información. (El primer enlace telefónico bidireccional por un solo tubo coaxial se tendió en 1957 entre Marsella y Argelia.)

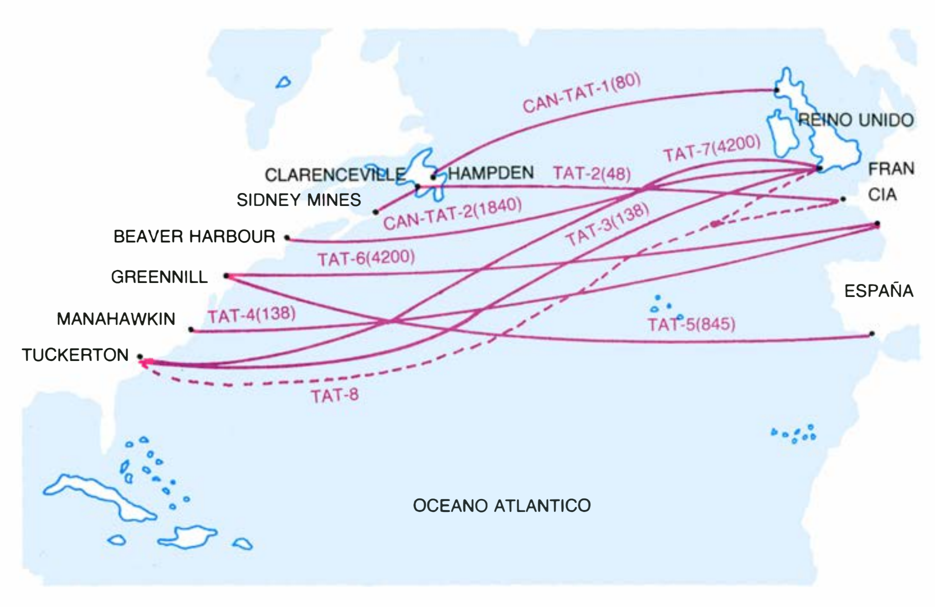
La tercera generación de cables submarinos viene marcada por el sello de las comunicaciones ópticas: en lugar de electrones se transmiten fotones y en vez de cables metálicos se emplean fibras de vidrio como soportes de la información. Gracias a esa innovación, se abaratan grandemente las comunicaciones submarinas, en parte debido a que el cable de fibra admite mayor volumen de comunicaciones simultáneas y en parte a las condiciones de la instalación. (Los cables de fibra óptica son más ligeros y precisan del orden de diez veces menos repetidores que los cables metálicos.)

La industria de los sistemas de comunicación por cable submarino se distribuye fundamentalmente entre el Reino Unido, Francia, Estados Unidos y Japón. A principios de nuestra década, la compañía británica STC producía alrededor del 38 por ciento del mercado mundial de sistemas por cable submarino; venían después la francesa Submarcom y la norteamericana Western Electric (esta última subsidiaria de la ATT), con un 25 por ciento cada una, aproximadamente. Y el 12 por ciento restante lo englobaba un consorcio de firmas japonesas integrado por Fujitsu, Nippon Electric y Ocean Cable Corporation. Las distintas empresas se repartían las zonas de actividad en razón de su entorno geográfico o histórico.

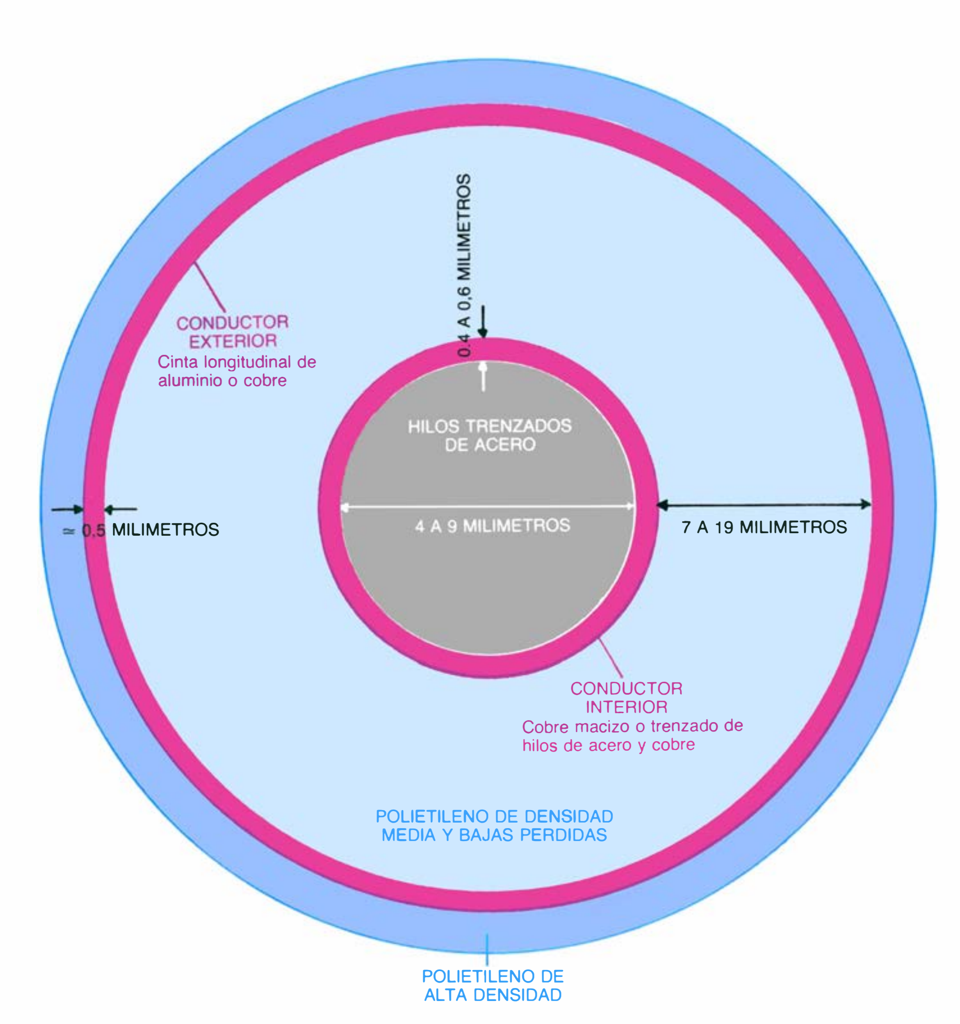
El volumen de ventas ha ido creciendo de una manera firme y regular: 695 millones de dólares estadounidenses se inviertieron durante el decenio 1956-1965, 956 millones durante 1966-1975, y 1192 millones entre 1976 y 1980. Estas inversiones hicieron posible la instalación de 54.642 millas de cable desde 1951 hasta 1965, que soportaban un total de 4200 canales telefónicos. De 1965 a 1976 se tendieron 53.035 millas de cable y 44.200 canales telefónicos; de 1976 a 1980, 31.719 millas de cable y 88.000 canales.

Adviértase, a tenor de esas cifras, que el coste por canal telefónico y milla ha bajado de 3,1 a 0,41 dólares USA desde 1955 hasta 1980, gracias, sobre todo, a la mejora de las técnicas de fabricación y al control de calidad de circuitos electrónicos y dispositivos, que permiten ya trabajar con frecuencias cercanas a los 100 megahertz. Merced a lo cual se puede aumentar el número de canales que se transmiten simultáneamente por un mismo cable. (A modo de comparación, un cable submarino instalado en 1950 soportaba apenas el centenar de canales telefó-

1. BARCO CABLERO. En sus bodegas puede albergar más de 12.000 milímetros cúbicos de cable, equivalente a unos 1500 kilómetros de cable-submarino convencional, constituido por conductores metálicos con estructura coaxial. Si se considera que un cable submarino de fibra óptica puede presentar una sección cinco o seis veces menor que los cables coaxiales, la capacidad de carga se multiplica y el tendido de los cables de fibra óptica puede hacerse con menos tiempo que el tendido de los cables metálicos convencionales.



2. PREDECESORES DEL PRIMER CABLE TRANSATLANTICO de fibra óptica. TAT-8 será el primer cable transatlántico de fibra óptica y entrará en servicio en 1988. Transportará 7500 canales telefónicos que podrán incrementarse hasta 37.800 mediante técnicas de multiplicación de circuitos. La capacidad de todos los cables submarinos existentes sobre el Atlántico no supera los 11.000 canales telefónicos. El TAT-7, uno de los más modernos, se instaló en 1983 y ofrece una capacidad de unos 4200 circuitos telefónicos.



3. CABLE COAXIAL SUBMARINO para aguas profundas. Cuando el tendido del cable ha de realizarse en la plataforma continental, deben añadirse varias armaduras adicionales para mejorar su resistencia a los agentes mecánicos. Para ofrecer adecuadas propidades mecánicas, los cables han de ser perfectamente simétricos, por eso los cables metálicos sólo incluyen un par de conductores en configuración coaxial.

nicos, en tanto que en 1980 se instalaban o proyectaban cables para soportar del orden de 4000 canales.)

A lo largo de los últimos 30 años, la demanda de cables telefónicos transoceánicos ha ido creciendo a un ritmo casi constante. En el Atlántico Norte, donde mayor es la demanda, se ha instalado un cable nuevo cada cuatro o cinco años; en otras palabras, el crecimiento anual en circuitos telefónicos fue del 20 al 27 por ciento. El número de circuitos telefónicos es siempre bastante menor que el número de conferencias establecidas. Un circuito físico de un cable submarino viene a cursar más de mil conferencias anuales. En cualquier caso, el crecimiento relativo del número de circuitos físicos instalados no debe ser inferior al crecimiento relativo del número de conferencias. En concreto, y centrándonos de nuevo en el Atlántico Norte, en 1983 se instaló entre Europa y Estados Unidos el cable TAT-7 para 4200 canales telefónicos; en 1988 está previsto que entre en servicio el TAT-8 (primer cable transatlántico de fibra óptica) con unos 7500 circuitos físicos.

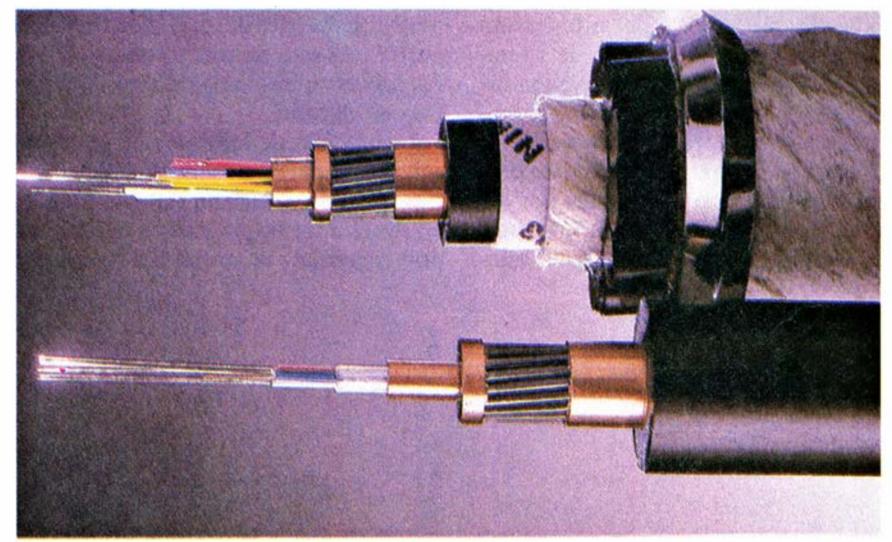
El coste del cable submarino frente al satélite depende del volumen de tráfico y del grado de proximidad geográfica entre los puntos a interconectar: los cables son más ventajosos en distancias cortas, en particular si, además, el volumen de tráfico es alto. Importa considerar también el retardo: una conversación por cable tarda 65 milésimas de segundo entre Europa y Canadá (demora que no advierten los interlocutores), mientras que por satélite el retardo se multiplica por 10 (alrededor de medio segundo). La propia calidad de la conversación es mejor a través del cable. Sin olvidar, por último, la vida media: cifrada en 20 años la del cable submarino y, en siete, la del satélite.

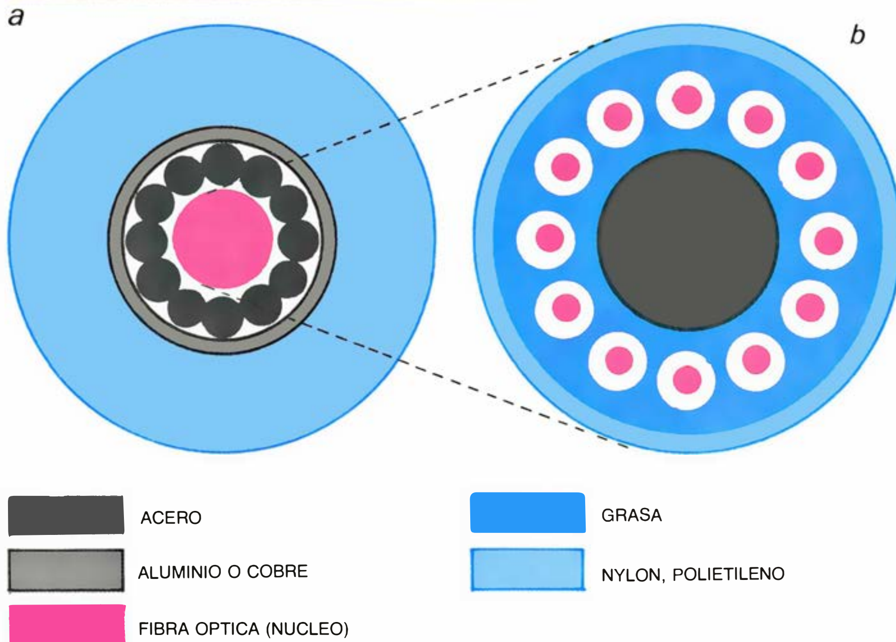
A finales de 1985 existían un total de 11.000 circuitos telefónicos instalados en cable submarino y unos 20.000 circuitos en uso regular por satélite. El cable TAT-8 añadirá otros 7500 circuitos telefónicos, los cuales pueden ascender hasta 37.800 empleando técnicas de "multiplicación" de circuitos, que consisten en aprovechar los espacios de tiempo muerto que hay en una conversación telefónica normal, para canalizar por un mismo circuito telefónico físico dos o más conversaciones telefónicas simultáneas. También a últimos del año pasado entraron en servicio nuevos satélites de la serie *Intelsat VI*, lo que permitirá disponer de más de 30.000 circuitos telefónicos transatlán-

ticos vía satélite. Estas cifras muestran el equilibrio en la repartición de comunicaciones entre satélites y cables submarinos.

La tecnología de las comunicaciones por fibra óptica, particularmente en los últimos cinco años, avanza en dos frentes: en las fibras que actúan como portadores de luz y en los elementos electroópticos asociados a los extremos del cable y en los puntos de regeneración. En estos últimos, los láseres transvasan a fotones la información que les llega en forma de señales eléctricas y los fotodetectores, de avalancha o tipo PIN, restauran al formato eléctrico la información que les entrega, en forma de luz, la fibra. Gracias a lo cual las fibras ópticas sustituyen con ventaja a los cables metálicos en los sistemas submarinos. Pero, ¿cuáles son, en definitiva, las razones que justifican el paso hacia los cables de fibra óptica, lo que antes llamábamos tercera época de los cables submarinos?

Primera: las fibras ópticas de vidrio de alta calidad, como las empleadas en los cables submarinos, tienen un diámetro de 125 micras (una micra es una milésima de milímetro). Son de dióxido de silicio. Aunque los cables de fibra incluyen también ciertos elementos, algunos metálicos, para proteger y dar consistencia al cable, tanto el peso de éste como su diámetro exterior es inferior al de los cables submarinos que emplean conductores metálicos para transmitir la información. A efectos comparativos, la sección transversal del cable coaxial del sistema TAT-6 (tendido en 1976 entre Europa y Norteamérica) es unas seis veces superior al de varias estructuras equivalentes de cable submarino por fibra óptica propuestas por fabricantes europeos, japoneses y estadounidenses. Ello implica que los buques cableros podrían ubicar en sus bodegas una longitud de cable de fibra seis veces superior que de cable convencional de tipo coaxial. Un buque cablero alberga unos 1500 kilómetros de cable de 5 centímetros de diámetro, equivalente a unos 12.000 metros cúbicos. Ello repercute en un menor coste y tiempo para la instalación del cable. No obstante, la capacidad de carga de cable viene limitada también por el peso: actualmente, los buques cableros suelen cargar un máximo de ocho a diez mil toneladas de cable. Los cables de fibra pesan siempre menos que los cables coaxiales metálicos, pues si bien en ambos casos la mayor parte del peso corresponde a los elementos de refuerzo para dar solidez al cable, los de fibra no incluyen el conductor exterior de los cables coaxiales, conductor que representa un 20 por ciento del peso, en los tipos de cable que se usan en aguas profundas.



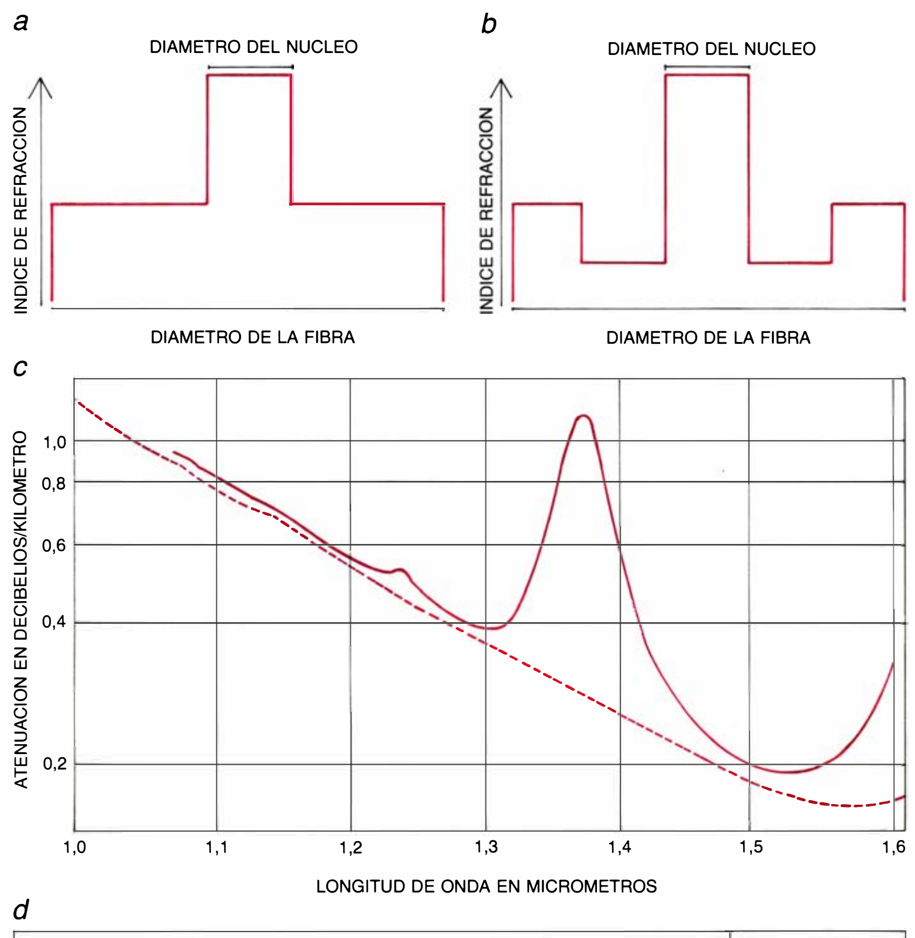


4. CABLES SUBMARINOS DE FIBRAS OPTICAS para aguas profundas y plataforma continental. Cuando se tienden en aguas menos profundas se agrega armaduras de hilos de acero para protegerlos frente a las corrientes marinas y redes pesqueras de arrastre. El cable consiste en un núcleo central que suele incluir de 4 a 12 fibras ópticas. Este núcleo se rodea con un cilindro metálico para protegerlo frente a la presión del agua. Por ese mismo cilindro puede enviarse la corriente eléctrica que necesitan los repetidores sumergidos. Abajo, a la izquierda, se ilustra un esquema de un cable submarino de fibra óptica para aguas profundas (*a*); a la derecha, una imagen de cables de fibra óptica para aguas profundas y plataforma continental (*b*).

Segunda ventaja: el producto "número de canales por separación entre repetidores". Importa que este producto sea alto, pues influye directamente en el coste por canal y kilómetro de cable. Las fibras ópticas ofrecen dos peculiaridades del mayor interés. Por una parte, si la fibra de vidrio se fabrica con los debidos requisitos de pureza y diseño, presenta muy baja atenuación a la luz que por ella viaja; por otra, algunos tipos de fibras ópticas (en particular las llamadas monomodo) ofrecen una anchura de banda mucho más elevada que los cables metálicos. La anchura de banda es inversamente proporcional al ensanchamiento que sufren los impulsos de luz al viajar por la fibra, y como la información se codifica en forma de impulsos de luz, cuanto mayor sea la anchura de banda del medio de transmisión mayor será el número de impulsos que se pueden enviar, sin que se interfieran entre sí, por unidad de tiempo. O lo que es lo mismo, cuanto mayor sea la anchura de

banda de un medio de transmisión tanto mayor será el número de comunicaciones simultáneas que pueden establecerse a su través. En síntesis, reuniendo estas dos propiedades de las fibras ópticas, es posible alcanzar productos de “número de canales por separación entre regeneradores” de unos 250.000 canales telefónicos por kilómetro. Esta cifra es unas diez veces superior a la que se puede conseguir con los cables submarinos de tipo metálico más modernos.

Los canales de información se expresan en unidades de información binarias. En este orden, un canal telefónico equivale a un flujo de unos 64.000 bits por segundo; un canal de televisión equivale a unos 18^8 bits por segundo; un canal de telegrafía a unos 50 bits por segundo, etcétera. El producto “canales por kilómetro” se establece en bits por segundo y multiplicado por kilómetro. De este modo hay menos ambigüedad y por un mismo sistema podrá transmitir un número de canales distintos en función de cuál sea la velocidad de transmisión (en bits por segundo) que requiera cada servicio de información. A tenor de estas consideraciones, podemos decir que el producto “velocidad de transmisión por separación entre puntos de regeneración” es de unos 3×10^{10} bit/s × km en los actuales sistemas de transmisión por fibra óptica. Valor que se multiplicará por cuatro o cinco en los próximos diez años.



d

ATENUACION (INCLUIDO EMPALMES) A 1300 NANOMETROS	≤ 0,5 dB/km
INCREMENTO DE ATENUACION AL CABLEAR LA FIBRA	≤ 0,02 dB/km
INCREMENTO DE ATENUACION A 1300 NANOMETROS POR EFECTOS AMBIENTALES (TEMPERATURA, PRESION, TENSION)	≤ 0,05 dB/km
RESISTENCVIA A LA TRACCION MECANICA	≃ 1,4 GPa (2 % ELONGACION)
VARIACION MEDIA DE LA ATENUACIÓN ENTRE 1290 Y 1330 NANOMETROS Y RESPECTO AL VALOR MINIMO	≤ 5 %
RANGO DE LONGITUDES DE ONDA DE DISPERSION MINIMA	1310-1320 nm
VALOR MEDIO DEL INCREMENTO DE ATENUACION POR PENETRACION DEL HIDROGENO	≤ 0,01 dB/km (EN 25 AÑOS)
VALOR MEDIO DEL INCREMENTO DE ATENUACION POR EFECTO DE LAS RADIACIONES EN EL FONDO MARINO	≤ 0,008 dB/km (EN 25 AÑOS)
DISPERSION ENTRE 1290 Y 1330 NANOMETROS	≤ 2,5 ps/km.nm

5. FIBRAS OPTICAS PARA CABLES SUBMARINOS. La tecnología actual ofrece dos grandes opciones, dentro de las llamadas fibras monomodo: fibras con perfil de índice de refracción en forma de escalón y fibras que presentan un revestimiento hundido. En *c*) se indica la forma de la atenuación, en función de la longitud de onda, obtenida con fibras de revestimiento hundido. En *d*) se reflejan las características principales del estado de la tecnología para fibras monomodo de aplicaciones submarinas en 1,3 micrometros.

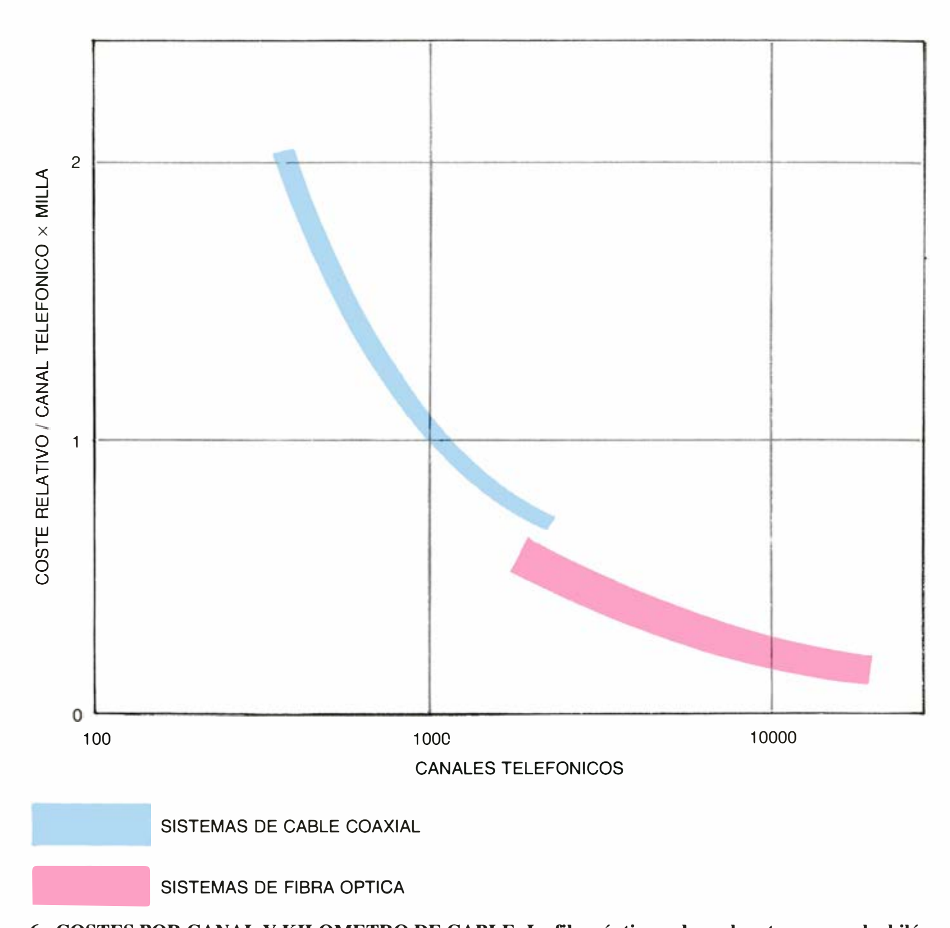
Vayamos con la última ventaja: para cumplir los altísimos grados de resistencia mecánica, los cables submarinos actuales se construyen manteniendo una configuración axial simétrica. Portan así un solo tubo coaxial, en tanto que los de fibra pueden incluir varios pares (seis o más fibras). Esto hace que la capacidad de información sea mucho mayor en los cables de fibra que en los conductores metálicos coaxiales. (En los cables metálicos no se han superado los 10.000 circuitos por cable; los de fibra podrán en breve soportar 100.000 circuitos por cable.) Esto abaratará el coste por circuito y kilómetro. Por su pequeño tamaño, un cable submarino puede llevar varias parejas de fibras y, por ende, ampliar su capacidad de transmisión; posibilita, además, la bifurcación de un cable en varias direcciones, incluso en medio del océano, salvando así la exigencia de limitarse a enlaces punto a punto, que es propia de los cables submarinos de conductores metálicos.

La diferencia básica entre un sistema de fibra óptica para aplicaciones submarinas o aplicaciones terrestre estriba en la fiabilidad que se pide a los cables submarinos en general y a los de fibra óptica en particular. Piénsese que una avería situada en un punto del fondo marino implica unos elevados costes de reparación y una pérdida de ingresos, mientras esté cortado el servicio por el cable. En los cables terrestres, las averías resultan siempre más cómodas de reparar y mientras dura el corte del cable se canaliza el tráfico por otras vías alternativas.

La fiabilidad se especifica, estadísticamente, en términos del llamado tiempo medio entre fallos (TMEF). Para los sistemas de cable submarino se ha

venido exigiendo un TMEF, del sistema entero, de 8 a 15 años. Para el primer cable transatlántico por fibra óptica se exige que no debe sufrir más de tres reparaciones en sus 25 años de vida. Estos requisitos globales se traducen en que los componentes eléctricos, electrónicos, ópticos y electroópticos deban tener una fiabilidad de unos pocos FITS (fallos en 10^9 componentes por hora). Cada componente se somete a un programa de cualificación en su fase de desarrollo; después, durante su fabricación, todos los componentes se inspeccionan, se someten a pruebas de envejecimiento (cuyo fin es detectar los defectuosos y no dañar los componentes correctos) y se prueban. Todo este proceso no persigue otra cosa que cada componente de un sistema de cable submarino satisfaga, en el entorno submarino a que estará sometido, los objetivos de fiabilidad.

La redundancia también resulta necesaria para satisfacer los objetivos globales de fiabilidad y mantener al mismo tiempo unos costos razonables. Hay al menos dos formas: una es la llamada redundancia activa (que se emplea con los diodos láser, por ejemplo) en que se conectan dos o más elementos de modo que el bloque en su conjunto funcione hasta que fallan todos sus elementos; y la llamada redundancia colateral ("stand-by"), en la cual uno o más elementos se mantienen en una reserva inoperante y sólo entran en servicio cuando falla el componente o elemento que estaba en estado operativo.

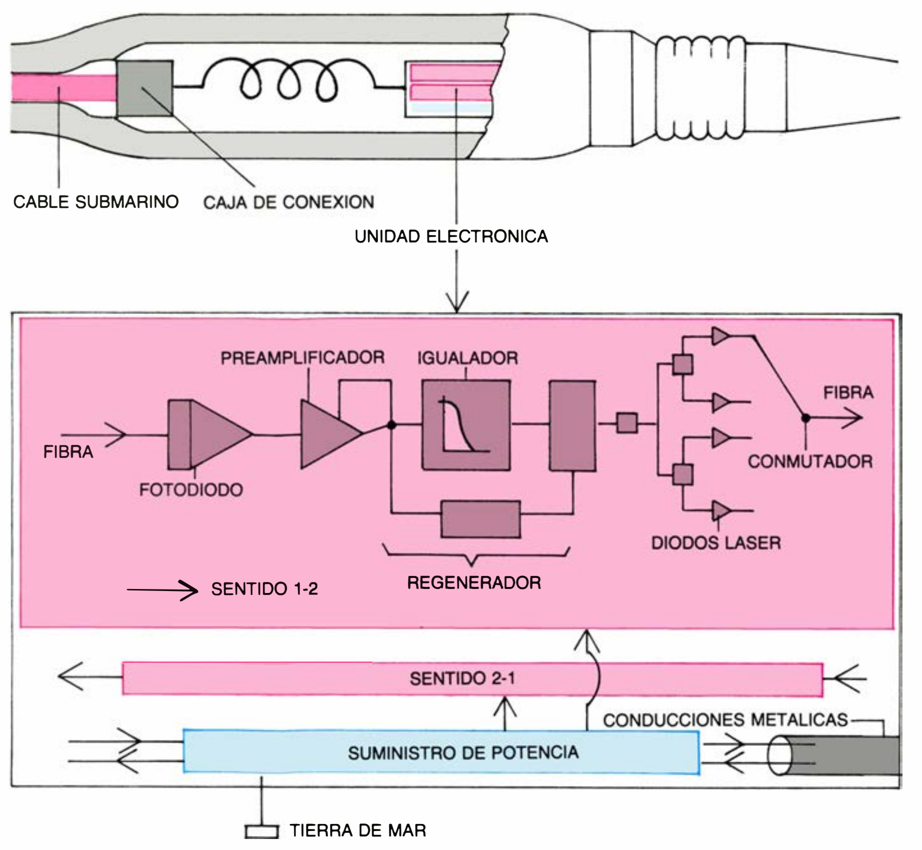


6. COSTES POR CANAL Y KILOMETRO DE CABLE. La fibra óptica reduce el coste por canal y kilómetro, que ofrecen los actuales cables submarinos metálicos, entre otras, por dos razones: *a*) la fibra óptica permite distanciar alrededor de diez veces más que los cables metálicos la separación entre repetidores; y *b*) los cables metálicos sólo incluyen un par de conductores coaxiales, mientras que los cables de fibra óptica suelen incluir, si lo comparamos con el de 4 a 12 fibras. En consecuencia, el número de canales transmitidos es mayor en los cables de fibra óptica que en cables convencionales de conductores metálicos coaxiales.

De acuerdo con la experiencia acumulada durante más de 30 años, se acepta distribuir las exigencias de fiabilidad al 50 por ciento entre el cable y los equipos electrónicos (fundamentalmente los repetidores submarinos). En relación a estos últimos, desempeña un papel importante la distancia entre repetidores, ya que cuanto mayor sea ésta menor será el número de repetidores y menos estrictos los requisitos de fiabilidad por repetidor. Veamos un ejemplo: el sistema submarino por cable metálico TAT-6, entre Europa y Norteamérica, con 5,1 kilómetros entre repetidores, presenta unos objetivos de fiabilidad de unos 20 FITS por repetidor. Un cable de fibra óptica con una separación de unos 30 km entre repetidores podría necesitar una fiabilidad de 120 FITS por repetidor (seis veces más baja que para el TAT-6).

La fiabilidad de un repetidor submarino para fibra óptica viene condicionada esencialmente por los fotodiodos láser y los fotodetectores. Al objeto de maximizar la separación entre repetidores interesa que los emisores y detectores de luz trabajen en longitudes de onda de 1,3 o 1,55 micrometros. A esas longitudes de onda, la atenuación de las fibras pasa por unos mínimos, y, además, su anchura de banda se maximiza, ya que uno de los mecanismos que la degradan –la dispersión temporal cromática– se hace muy pequeña. Las fibras que presentan mayor anchura de banda son las llamadas de tipo monomodo. Consisten en un cilindro de vidrio de 125 micras de diámetro con un núcleo de 8 a 10 micras por el cual viaja la luz, y son las más idóneas para los cables submarinos. Los ya existentes sistemas terrestres por cable de fibra monomodo ofrecen, en condiciones normales de explotación, productos de hasta 4×10^{10} bit/s × km, equivalentes a transmitir unos 14.000 canales telefónicos con una distancia entre repetidores de 40 km.

La separación entre los repetidores no depende sólo de la fibra, sino también del margen de atenuación que se pueda admitir en la señal óptica desde que sale del fotoemisor láser hasta que llega al fotodetector (un fotodiodo de avalancha o un fotodiodo PIN) del siguiente repetidor. Veamos un ejemplo: un láser para fibras monomodo puede inyectar del orden de −2 dBm de potencia óptica en la fibra, y un fotodetector requiere del orden de −40 dBm o −36 dBm para garantizar una tasa de de 10^{-9} errores a una velocidad de 274 megabit por segundo. En consecuencia, si el cable óptico atenúa del orden de 0,8 dB/km, los repetidores podrían separarse, como máximo, de 42 a 47 kilómetros. (Una magnitud positiva se expresa en *dB* aplicándole el logaritmo en base diez y multiplicando por diez el resultado. El número de *dBm* que tiene una cantidad de potencia es diez veces el logaritmo en base diez de esa potencia expresada en megawatt.) También debe considerarse que la pérdida por kilómetro no depende sólo de la fibra óptica en sí, sino también del proceso de cableado (protección e inclusión de las fibras dentro de la estructura del cable), así como de los empalmes. Según ensayos realizados por la Compañía Telefónica Nacional de España, se ha visto viable obtener empalmes con ate-



7. REPETIDOR SUBMARINO PARA UN CABLE OPTICO. Corresponde a un esquema genérico y representa una unidad electrónica bidireccional; incluye también los empalmes entre las fibras del cable y las fibras asociadas a las unidades electroópticas que realizan la amplificación y regeneración. Actualmente, la regeneración se lleva a cabo a nivel eléctrico: convirtiendo la señal óptica en señal eléctrica, tratando la señal eléctrica y volviendo a realizar la reconversión óptica. Es usual que los repetidores incluyan varios diodos láser, de los que sólo uno está activado; se conmutan automáticamente cuando el láser activo se avería.

nuaciones inferiores a 0,1 o 0,2 dB. Los empalmes han de hacerse para intercalar los repetidores en el cable y para juntar las longitudes unitarias de fabricación del cable.

Además de estos requisitos de transmisión, por influir directamente en la calidad de las señales electrónicas o de las señales ópticas, los cables submarinos deben satisfacer unos requisitos mecánicos acordes con las condiciones del tendido del cable y del lecho submarino. En particular, las fibras ópticas han de soportar importantes fatigas a la tracción, no sólo en la instalación sino ante una eventual recuperación del cable desde su ubicación en el fondo marino. Para garantizar el buen comportamiento de las fibras ante los esfuerzos mecánicos, se las somete, después de su fabricación, a diversas pruebas de tracción simulando los esfuerzos a que se hallarán durante la instalación, condiciones normales de operación o bien una eventual recuperación del cable. Esta última operación es la más exigente y requiere someterlas a pruebas de tracción equivalentes a una elongación del 2 por ciento: la longitud unitaria de los largos de fabricación del cable óptico depende de cuáles sean las longitudes de fibra que soportan sin fracturarse una elongación del 2 por ciento.

En España, entre las islas de Tenerife y Gran Canaria, se instaló a finales de 1985 el primer cable comercial submarino en aguas profundas por fibra óptica incluyendo repetidores, fruto de un acuerdo de colaboración entre la Compañía Telefónica y la American Telephone and Telegraph. La distancia entre ambas islas es de unos 117 kilómetros de cable, tendido a profundidades de hasta 3000 metros. Se instalaron tres puntos intermedios con repetidores, alejados entre sí unos 30 kilómetros. El cable incluye seis fibras ópticas; de los canales, dos se reservarán para realizar experimentos relativos a la distribución estadística de los errores de transmisión, la acumulación del parpadeo de la señal digital causada por los equipos electrónicos que regeneran la señal en los puntos de repetición, la influencia de la temperatura y del ambiente marino en el cable y en los equipos electrónicos y electroópticos, etcétera. Cada una de las dos parejas de fibras destinadas a uso comercial pueden soportar 3840 canales telefónicos físicos.

Las fibras del cable entre Tenerife y Gran Canaria están diseñadas con un perfil de índice de refracción en forma de un núcleo central de 8,3 micras y un revestimiento de bajo índice de refracción con un diámetro de 52,3 micras. El resto, hasta formar un cilindro de 125 micras, es de óxido de silicio. Cada una de las fibras del cable submarino ha sufrido unas pruebas de resistencia a la tracción que equivalen a una elongación del 2 por ciento. Ya en el cable, las fibras presentan pérdidas inferiores a 0,49 dB/km a 1300 nanómetros.

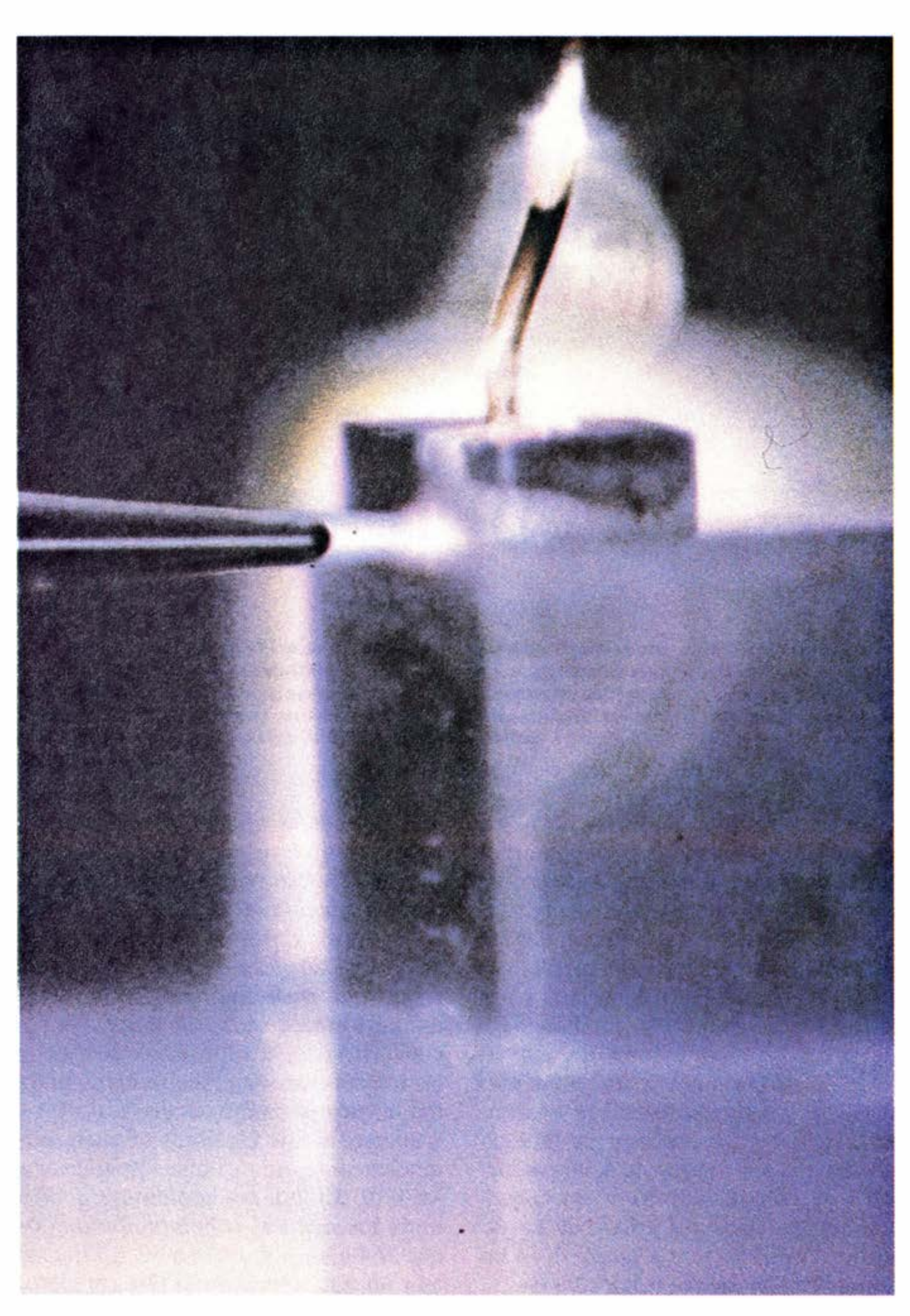
En Estados Unidos las actividades de investigación y desarrollo de cables submarinos por fibra óptica están coordinadas por la American Telephone and Telegraph. En 1981 se llevó a cabo con éxito la demostración en laboratorio de un sistema que transportó 3600 canales telefónicos por dos fibras. Al año siguiente, se realizó un ensayo en aguas profundas que demostró la viabilidad de la tecnología de fibra óptica. Este ensayo dio lugar a un rediseño final con los equipos instalados, en 1985, en las Canarias. Equipos similares serán los que se instalen en el cable transatlántico TAT-8, entre la costa de Estados Unidos y un punto cercano a Europa, donde el cable se bifurca hacia Francia y hacia el Reino Unido.

En Japón, la Kokusai Denshin Denwa Company (KDD) y la Nippon Telegraph and Telephone Corporation (NTT) explotan, en régimen de monopolio, las telecomunicaciones con el exterior y las nacionales, respectivamente. La KKD comenzó los estudios de laboratorio en 1976; y en 1979 ya disponía de un sistema prototipo, capaz de operar en aguas profundas, que soportaba 3780 circuitos por un par de fibras ópticas. En 1982 y 1984 realizó ensayos con prototipos mejorados de este sistema, bautizado con el nombre de os-280. La KDD ha anticipado el uso de fibras ópticas para el enlace submarino previsto para 1988 entre Japón y Hawai. Por su parte, la NTT introdujo en 1985 los primeros enlaces submarinos de ámbito nacional por fibra (enlaces sin repetidores).

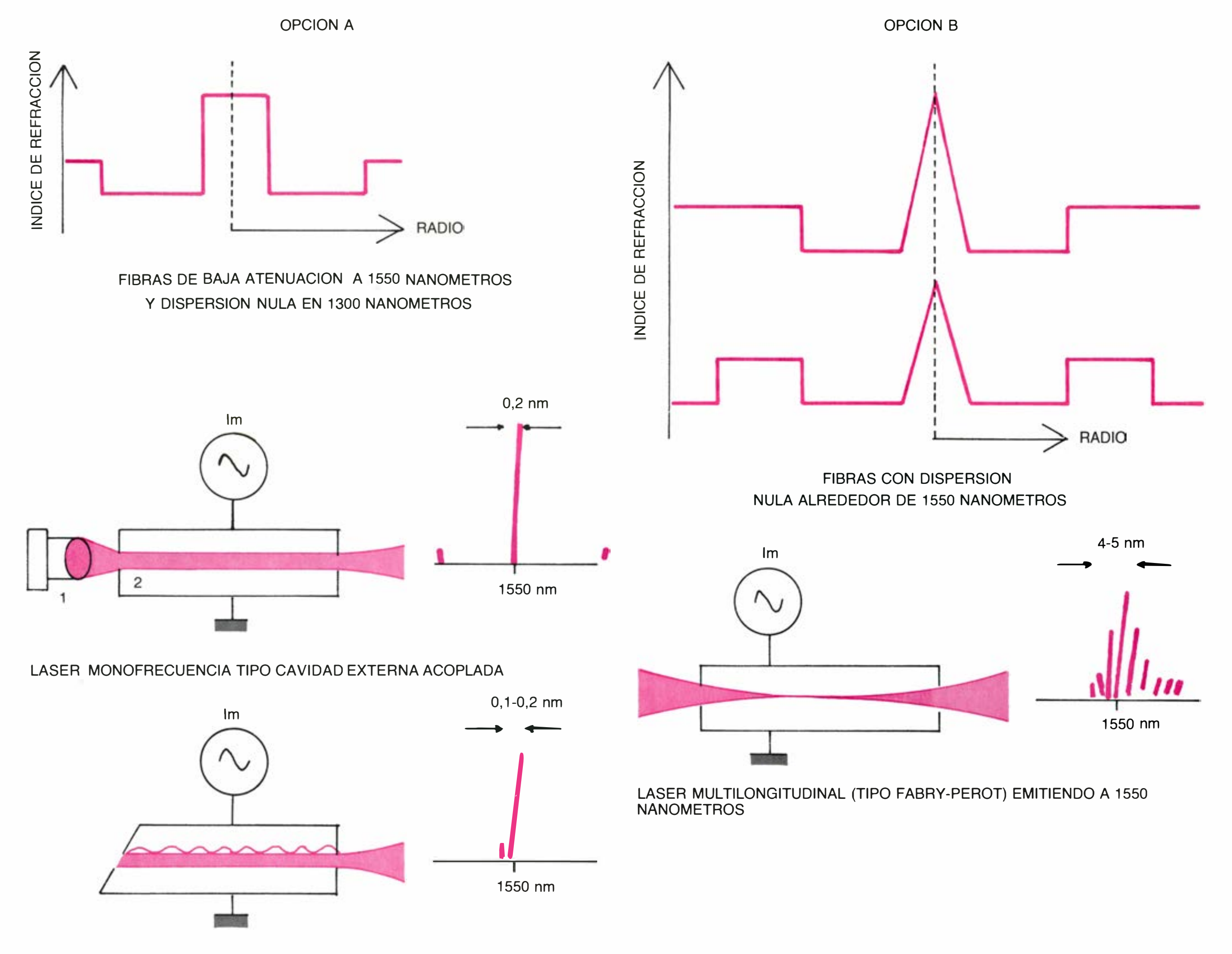
En el Reino Unido las actividades de investigación y desarrollo sobre cables submarinos de fibra óptica se centran en la Standard Telephone and Cables (STC) en cooperación con los laboratorios del British Telecom en Martlesham. Fruto de este esfuerzo es el desarrollo de un sistema denominado NL-2, entre el Reino Unido (Broadstairs) y Bélgica (Ostende), que entraría

en servicio comercial en este año de 1986. Una nueva versión del actual NL-2 se instalará en 1988 para cubrir los 500 kilómetros que enlacen el TAT-8 con la costa del Reino Unido (en Widemonth). Este sistema permite enviar 3840 canales telefónicos por cada dos fibras; su objetivo de fiabilidad es de un máximo de tres fallos durante sus 25 años de vida. Este sistema NL-2 es una nueva versión de otro anterior, NL-1, desarrollado por la propia STC en colaboración con el British Telecom. Como antecedente del NL-1, la compañía STC instaló en 1980 el primer cable de fibra óptica del mundo, con una longitud de 9,5 kilómetros en Loch Fyne, Escocia.

A finales de los años 70, la industria francesa (principalmente CIT-Alcatel y Cables de Lyon) en colaboración con la Administración de Correos y Telecomunicaciones, acordó poner en marcha un programa de cinco años, para desarrollar un sistema de cable óptico submarino, denominado S-280. Este sistema es capaz de transmitir 3840 canales telefónicos por un par de fibras ópticas. El proyecto contempla que el cable puede instalarse hasta 6500 metros de profundidad con una separación entre los repetidores no inferior a 45 ki-



8. DIODOS LASER. Constituyen una pieza clave en la viabilidad de los sistemas de cable submarino por fibra óptica. Los que se están empleando en la primera generación de cables submarinos pueden alcanzar velocidades de transmisión de varios centenares de megabit por segundo, ofreciendo potencias del orden de megawatt. Emiten en longitudes de onda alrededor de los 1300 nanómtros y aunque producen un modo transversal único, la radiación luminosa la distribuye entre modos de varias longitudes de onda.



9. OPCIONES TECNOLOGICAS FUTURAS en cables submarinos de fibra óptica. El número y la diversidad de mejoras y opciones se multiplican día a día. Ahora bien, el período de tiempo entre la demostración de la viabilidad tecnológica de una determinada opción y su realización práctica en aplicaciones comerciales oscila entre siete y diez años. Por ello, si pensamos en una segunda generación de cables ópticos submarinos hacia 1992 o 1993, las tecnologías básicas de esos sistemas deberían estar ya suficientemente demostradas. Por ello, en opinión del autor del artículo, esa próxima generación habrá de utilizar láseres emitiendo alrededor de 1550 nanómetros, bien con las fibras parecidas a las actuales o bien de núcleo con perfil de índice triangular o segmentado.

lómetros. Para mejorar la fiabilidad del sistema, cada repetidor incluye tres láseres de reserva por cada láser activo.

Recientemente, en abril de 1984, se hizo el primer ensayo de cable submarino con repetidores; el cable se instaló en la Rivera francesa, con 80 kilómetros de longitud y una profundidad máxima de 1800 metros, y los repetidores se ubicaron a 1000 metros de profundidad. Fruto de este período de trabajo es la instalación de un primer enlace comercial por cable submarino de fibra óptica entre Córcega y un punto de la costa continental, ubicada a unos 400 kilómetros. Una posterior versión de este sistema será el que cubra en 1988 los últimos 330 kilómetros para enlazar con Francia (en Ponmarch) al TAT-8.

En un horizonte de 8 a 10 años, los cables submarinos podrán emplear diodos láseres transmitiendo alrededor de 1,55 micras, ya que la atenuación de las fibras es menor a esas longitudes de onda; de esta manera se reduce el número de repetidores submarinos y, en consecuencia, mejora la fiabilidad. En los primeros años habrán de acometerse investigaciones para aumentar nuestros conocimientos sobre el comportamiento y los procesos de fabricación de diodos láseres de semiconductores de compuestos III-V; los resultados serán decisivos a la hora de escoger el tipo de fibra a emplear en las futuras transmisiones a 1,55 micras.

Hay dos opciones básicas: seguir con las fibras actuales usadas en esta primera generación de cables submarinos, o bien acudir a nuevos diseños con perfiles de índices triangulares o segmentados. También están emergiendo otras tecnologías que podrían ser candidatas para la generación siguiente de cables submarinos. Entre éstas podríamos señalar: las técnicas de detección coherente (que posibilitan amplificar la señal en forma de luz, sin tener que reconvertirla antes a formato eléctrico) y las transmisiones en longitudes de onda del infrarrojo medio, es decir, de 1,4 a 4 micras, donde las fibras ofrecen, potencialmente, una atenuación inferior a los 0,01 dB/km. No obstante, en 1986 tanto los sistemas coherentes como los que trabajan por encima de 1,6 micras aún no han demostrado por completo su viabilidad; por ello no cabe esperar que estas tecnologías se incorporen a los sistemas de cables submarinos antes de 1995.

Dinámica de las proteínas

Las moléculas fundamentales de la vida nunca están en reposo; si fueran rígidas, no podrían desempeñar sus funciones. Mediante simulaciones por ordenador, se han explorado los incesantes movimientos que subyacen a su actividad

Martin Karplus y J. Andrew McCammon

El estudio de cómo las proteínas satisfacen las necesidades de los organismos vivos constituye un curioso ejemplo de metodología que, a la vez que permitió avances espectaculares, condujo a un error de concepto. El método al que nos referimos es la cristalografía por rayos X; en efecto, a partir de la difracción, o desviación, que sufre un haz de rayos X se determina la estructura de las moléculas de proteína contenidas en un cristal de la substancia analizada. La belleza intrínseca y el notable nivel de detalle de las estructuras resueltas por ese método indujeron a pensar en las proteínas como entidades rígidas. Lo que a su vez ha llevado a suponer, erróneamente, que los átomos de las proteínas ocupan posiciones fijas. La mayoría de los esfuerzos dirigidos a buscar una explicación del funcionamiento proteico (por ejemplo, la actividad enzimática, o catalizadora, de las proteínas) se han basado en estructuras estáticas. Creíase que la especificidad con la que una enzima se unía a una molécula particular, su sustrato, y la eficacia que mostraba al catalizar una reacción bioquímica derivaban de interacciones complementarias entre el sustrato y una enzima rígida, interacciones que se comparaban con el acoplamiento de una llave con su cerradura. Los casos en los que se sabía que la morfología (configuración) de la proteína cambiaba (pensemos en la hemoglobina enlazada con el oxígeno) solían reputarse transiciones bruscas entre estructuras por lo demás estáticas.

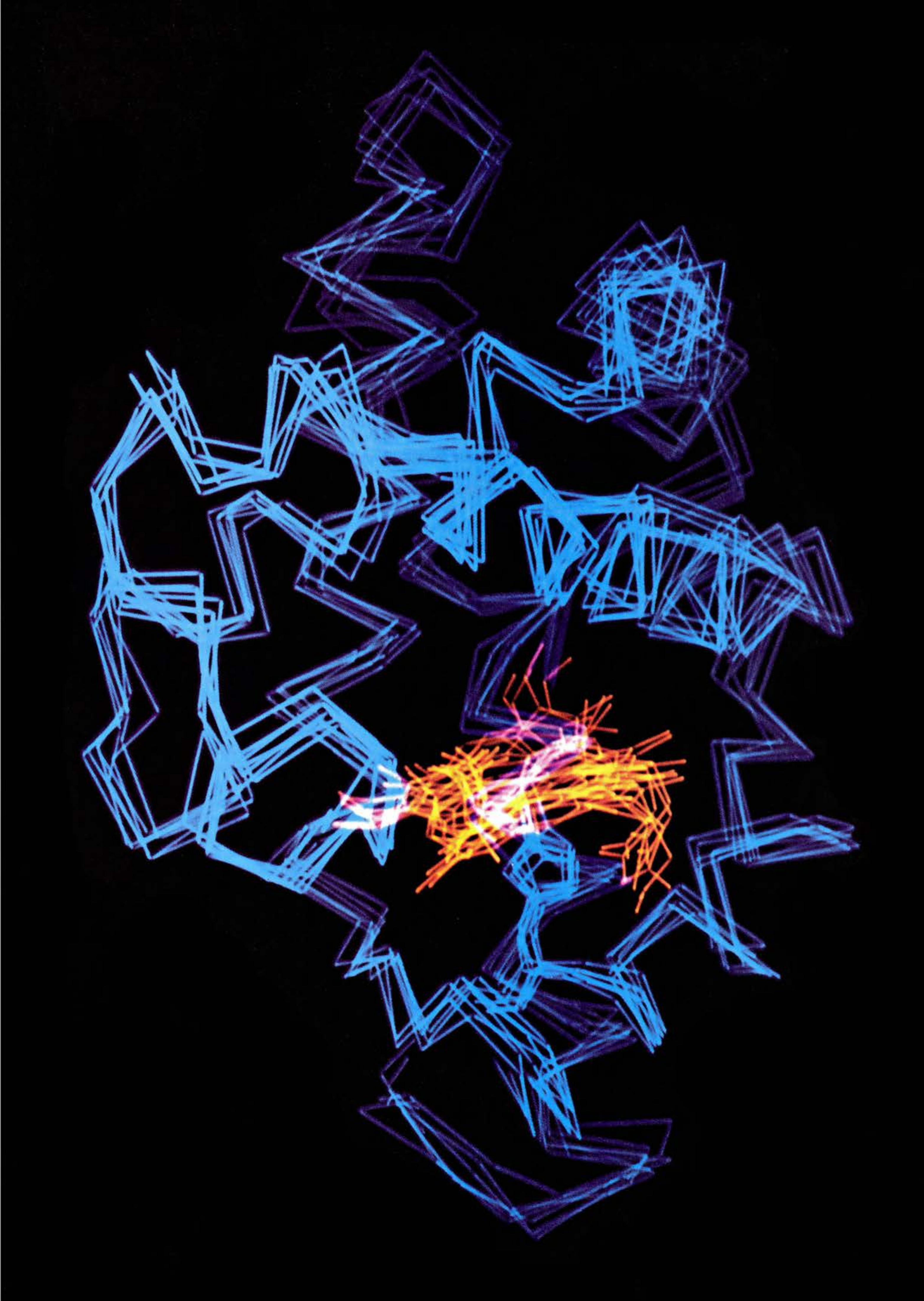
A lo largo de la última década se ha sometido a revisión esa imagen estática de las proteínas. Se reconoce hoy que los átomos de las moléculas de un polipéptido se encuentran en movimiento incesante, y que lo que el cristalógrafo observa constituye, a lo sumo, la estructura promedio de la proteína. Los átomos de toda molécula proteica experimentan fluctuaciones de alta frecuencia en torno a ese promedio. En pocas palabras, las proteínas son dinámicas: varían constantemente en su morfología. Un cristal listo para someterlo a análisis por rayos X viene a contener unas 10^{20} moléculas de proteína; sin embargo, resulta altamente improbable que, en un momento dado, ni siquiera una sola presente la estructura promedio.

No por ello resulta menos importante conocer las posiciones atómicas promedio; antes bien, sirven de punto de partida de cualquier intento de elaborar descripciones más ajustadas de la actividad de la proteína. La imagen dinámica permite entender fenómenos inexplicables con los modelos estáticos. Derivan esas descripciones dinámicas de muchas fuentes, en las que cabe destacar las técnicas experimentales, incluida la cristalografía por rayos X. Por su lado, los estudios teóricos han puesto de relieve detalles del movimiento proteínico cuyo descubrimiento difícilmente podría haberse esperado de las mediciones experimentales. En algunos casos, las simulaciones del movimiento atómico han demostrado que la actividad de la proteína resultaría imposible si ésta se encontrara fija en su estructura promedio.

La estructura y las fluctuaciones de las proteínas les permiten ejecutar muchas de las tareas que exige el funcionamiento de un organismo. Todas las enzimas son proteínas; esto es, las proteínas son los catalizadores que aceleran las reacciones esenciales de los sistemas vivos (incluida la propia síntesis de proteínas) para que procedan a la velocidad pertinente. Otras proteínas transportan moléculas de pequeñas dimensiones, electrones y energía hasta los puntos del organismo donde resulten precisos. A muchas proteínas les corresponde un papel estructural; por ejemplo, forman parte del tejido fibroso y del muscular.

Dada la variedad de sus funciones, sin duda existirán muchas proteínas distintas. Según se sabe, la bacteria *Escherichia coli*, organismo unicelular muy bien estudiado, contiene 3000 diferentes, algunas multiplicadas en gran número de copias; una sola bacteria puede poseer alrededor de un millón de proteínas. Todo polipéptido está formado por un número específico de pequeñas subunidades, los aminoácidos. Existen 20 aminoácidos distintos, caracterizados todos ellos por su cadena lateral: un grupo químico específico cuya complejidad va del átomo de hidrógeno, en el más sencillo de los aminoácidos, la glicina, a complejos anillos atómicos, en el más elaborado, el triptófano. En las proteínas, los aminoácidos aparecen unidos linealmente uno a continuación del otro: constitu-

1. MOVIMIENTOS DE UNA PROTEINA, registrados en una escala temporal de unidad el picosegundo (una billonésima de segundo). Se ilustran aquí superponiendo siete posiciones adoptadas por una molécula de mioglobina a intervalos de cinco picosegundos, según propone una simulación de la dinámica molecular efectuada por ordenador. Como cualquier proteína, la mioglobina es una cadena de aminoácidos (*azul*). Se muestra la cadena en una versión simplificada que sólo presenta la posición del átomo de carbono central de cada aminoácido. La cadena adopta una conformación plegada, tridimensional, que le es característica; en la mioglobina, esa conformación consta de ocho segmentos de hélice alfa conectados por pequeños bucles de unos pocos aminoácidos. En la simulación, las hélices se mueven sin perder su forma. La molécula de mioglobina contiene también una compleja molécula orgánica denominada grupo hemo (*naranja*), en cuya porción central se halla un átomo de hierro capaz de unirse al oxígeno; en efecto, la mioglobina almacena oxígeno en el tejido muscular. Esta representación por ordenador, al igual que las restantes que aparecen en el artículo, las generó John Kuriyan, de la Universidad de Harvard y el Instituto de Tecnología de Massachusetts, valiéndose del programa HYDRA, redactado por Roderick E. Hubbard, de la Universidad de York.

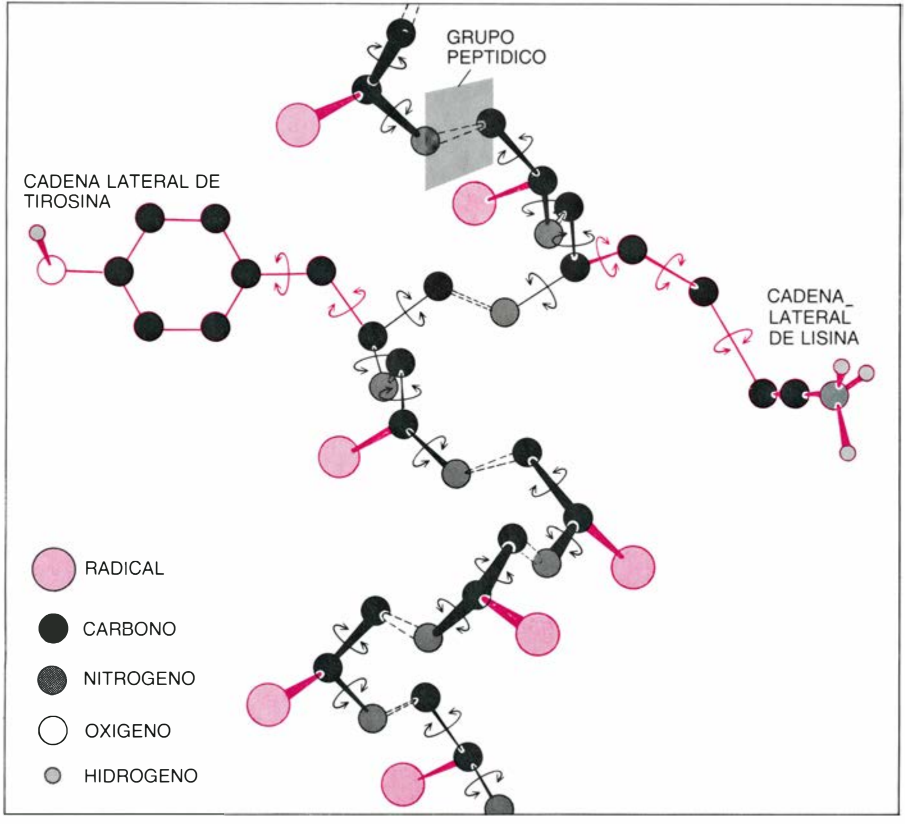


yen la cadena polipeptídica. Suelen contener esos polipéptidos de 50 a 500 aminoácidos, esto es, de 500 a 5000 átomos. La secuencia exacta que sigan los aminoácidos determina la estructura promedio, así como otras propiedades de la proteína. En particular, el equilibrio de las fuerzas de atracción y de repulsión entre los átomos que componen los aminoácidos provoca el repliegue de la cadena de las proteínas globulares, doblamiento que se demuestra esencial para sus funciones.

El primer paso a dar en la interpretación del papel que les corresponde a las fluctuaciones atómicas en la función proteínica es determinar la propia naturaleza de esas fluctuaciones. Se trata de analizar su magnitud (la extensión que abarcan), su probabilidad (la frecuencia con que se dan) y su escala temporal (cuánto duran). La manera más directa de abordar la dinámica de proteínas es considerar cada átomo como una partícula que responde a las fuerzas según prescribe la física newtoniana, esto es, de acuerdo con las ecuaciones del movimiento.

Hallamos una analogía sencilla en el cálculo del movimiento de un péndulo o de una partícula suspendida de un muelle. Dadas la posición y velocidad de la partícula, así como la magnitud y dirección de la fuerza que actúa sobre ella en un instante particular, por extrapolación, o integración, de las ecuaciones del movimiento se obtendrán la posición y la velocidad en un instante ligeramente posterior. Los resultados serán precisos si el intervalo de extrapolación, el denominado incremento de tiempo, es suficientemente breve para que la fuerza que actúa en la nueva posición no difiera mucho de la que se registraba en la original. Por iteración del procedimiento de extrapolación, esto es, integrando las ecuaciones newtonianas del movimiento en períodos temporales más prolongados, se obtiene la trayectoria de la partícula como una serie de posiciones y velocidades en tiempos sucesivos, una velocidad y una posición para cada incremento de tiempo.

Los enlaces químicos que se establecen entre los átomos de una cadena polipeptídica se comportan como si fueran muelles. No obstante, entre átomos no unidos por enlaces se registran también fuerzas de menor magnitud; por ejemplo, las que impiden que, en un mismo instante, ocupe más de un átomo la misma posición en el espacio. En una proteína que conste de muchos átomos, la fuerza total que actúe sobre cada uno de ellos en un momento dado dependerá, por tanto, de las posiciones que ocupen los demás. A nadie sorprenderá, pues, que la solución de las ecuaciones de Newton para el cálculo de las posiciones y velocidades de todos los átomos de una proteína obligue al empleo de ordenadores de alta velocidad. Constituyen esos cálculos lo que ha dado en denominarse simulación de la dinámica molecular.



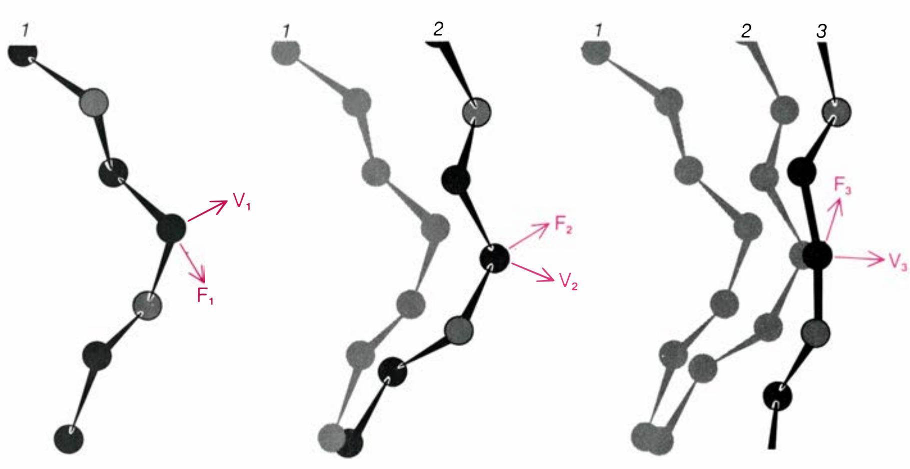
2. PUNTOS FLEXIBLES de una cadena polipeptídica, que le permiten plegarse y adoptar la conformación característica de la proteína de que se trate; por añadidura, facilitan la fluctuación de los átomos de la proteína con respecto a sus posiciones medias. Se representan aquí sólo los átomos principales de la cadena polipeptídica. El esqueleto de la cadena (*enlaces negros*) lo forman átomos de carbono y de nitrógeno; las uniones denominadas enlaces peptídicos son rígidas, mientras que las que quedan entre éstos permiten la rotación (*flechas de giro*). Del átomo central de cada aminoácido se proyecta una cadena lateral (*enlaces de color*); las cadenas laterales que se muestran en el dibujo con más detalle poseen también enlaces con capacidad de giro.

Para iniciar la simulación se requiere un conjunto representativo de posiciones de partida, que se deducen de los datos proporcionados por la cristalografía de rayos X. Las posiciones cristalográficas no sirven para configurar la estructura inicial, puesto que se trata de una estructura promediada: la localización de algunos átomos no se corresponde con ninguna estructura real (por ejemplo, puede que se sitúen dos átomos demasiado cerca). Tal distorsión estructural podría, a su vez, inducir a creer que sobre algunos átomos actúan fuerzas muy intensas, poco verosímiles.

El procedimiento denominado de equilibrado dinámico supera ese obstáculo y confiere a la proteína un estado tal que sus posiciones y velocidades resultan adecuadas para dar comienzo al cálculo de las trayectorias. En el procedimiento de equilibrado se ajustan las posiciones cristalográficas de los átomos de la proteínas (y los de todas las moléculas circundantes de solvente que quiera incluirse en la simulación), de modo que se relajen las fuerzas poco verosímiles. Se asignan entonces a los átomos pequeñas velocidades estocásticas correspondientes a cierta temperatura (próxima al cero absoluto). (La temperatura de una molécula constituye una medida de la magnitud de sus velocidades atómicas promedio.) A partir de ese conjunto de posiciones y velocidades, y de los valores de las fuerzas instantáneas, se permite el movimiento de los átomos de acuerdo con las ecuaciones newtonianas del movimiento, por ejemplo, unos 100 incrementos de tiempo. Se asignan a continuación conjuntos progresivamente superiores de velocidades aleatorias y se permite de nuevo el movimiento de los átomos, hasta alcanzarse la temperatura deseada (la temperatura ambiente, por ejemplo); se

habrá completado el equilibrado cuando el sistema evolucione espontáneamente durante cierto tiempo por integración de las ecuaciones del movimiento, manteniéndose estables la temperatura y la estructura medias.

La simulación de la dinámica molecular parte de la proteína equilibrada. Puesto que, en cada paso sucesivo de la simulación, las fuerzas que se ejercen entre sí los átomos sólo deben alterarse ligeramente, el incremento de tiempo debe ser breve (respecto de la escala temporal de los movimientos más rápidos de la proteína). Un incremento de tiempo típico vendría a ser del orden de 10^{-15} segundos (un femtosegundo). Se almacenan en cinta magnética, para posteriores análisis y representaciones gráficas, los cálculos de las posiciones y velocidades atómicas. En una simulación típica, las trayectorias resiguen la dinámica de la proteína a lo largo de un período de unos 10^{-10} segundos (100 picosegundos); los nuevos ordenadores, de velocidades superiores, permiten simular períodos de 10^{-9} segundos (un nanosegundo), o incluso más largos. (En este caso, que requiere hasta un millón de incrementos de tiempo, un superordenador debe dedicarle al cálculo varios centenares de horas.) Por fortuna, son numerosos los movimientos proteínicos de interés que se desarrollan en su totalidad en 100 picosegundos o menos. Incluso en tan escaso tiempo resulta prodigiosa la cantidad de detalle que se acumula en la trayectoria.



3. SIMULACION DE LA DINAMICA MOLECULAR. Trata los átomos que componen la proteína como si fueran partículas newtonianas. Se obtiene así una descripción detallada de sus movimientos. Los enlaces entre átomos se equiparan a muelles poco elásticos; los átomos no unidos por enlaces interactúan por medio de fuerzas más débiles. Las simulaciones se efectúan en una serie de breves incrementos de tiempo (de 10^{-15} segundos), en los que se calcula la velocidad y posición de cada átomo. Dadas las posiciones y velocidades en el tiempo t_1, y un método que determine las magnitudes y direcciones de las fuerzas que actúan sobre los átomos, las ecuaciones del movimiento darán las posiciones y velocidades al cabo de un incremento de tiempo (t_2). El cálculo de las fuerzas que se registran en t_2 permite determinar las posiciones y velocidades en t_3. Por iteración se conocen los movimientos internos de la proteína en intervalos de hasta 10^{-9} segundos.

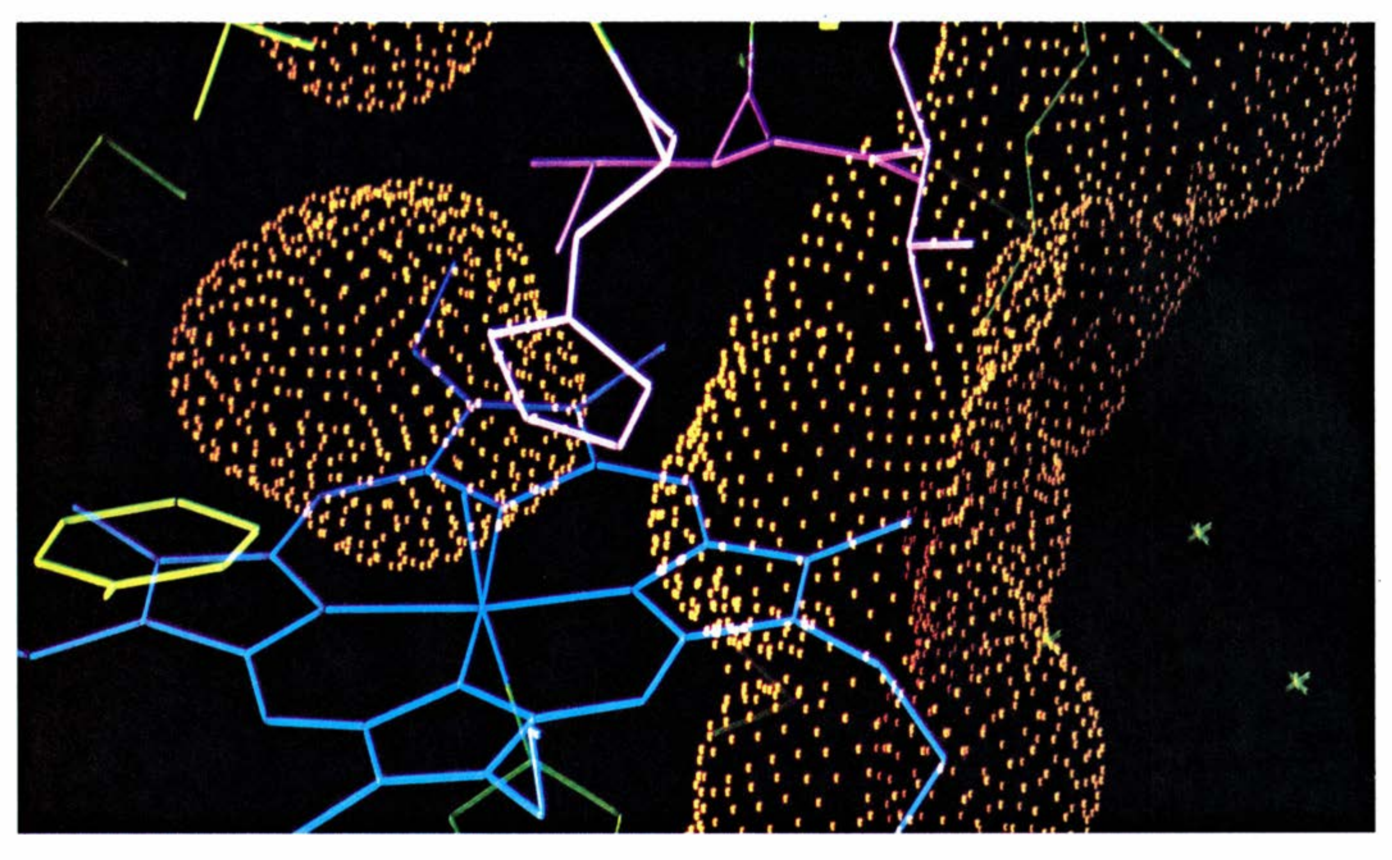
Los movimientos de los átomos de un polipéptido, revelados por las simulaciones de dinámica molecular, tienden a compartir ciertas características que reciben explicación en el marco de la estructura básica de las proteínas. Al igual que ocurre en la mayoría de polímeros, la cadena polipeptídica que constituye la proteína es flexible porque muchos enlaces fuertes que unen entre sí los átomos presentan libertad de rotación. La libertad de rotación del enlace permite que una porción de la cadena se mueva con respecto a la otra. En el giro de la cadena polipeptídica la acompañan sus cadenas laterales, que a su vez poseen también enlaces que permiten la rotación, lo que dota al conjunto de mayor flexibilidad.

Esa flexibilidad del esqueleto polipeptídico y de sus cadenas laterales confiere a las proteínas globulares su capacidad de plegarse en su estructura nativa, o promediada. Sin embargo, aun encontrándose plegada la proteína, la energía térmica correspondiente a las velocidades atómicas a temperatura ambiente basta para permitir los movimientos de rotación, que constituyen la fuente principal de fluctuaciones atómicas de la proteína. Las proteínas globulares, cuyos átomos se encuentran muy próximos unos de otros, ven restringidas sus fluctuaciones. Con todo, la suma de gran número de pequeños movimientos locales puede traducirse en desplazamientos a gran escala, globales, de una porción de la proteína respecto a otra.

En intervalos cortos, por ejemplo de una pocas décimas de picosegundo, las fluctuaciones más destacadas son movimientos locales en los que, por regla general, los grupos peptídicos y las cadenas laterales rotan entre 20 y 60 grados alrededor de los enlaces covalentes que ensamblan los grupos al resto de la cadena polipeptídica. Suelen esos movimientos presentar poca coherencia. Los grupos en rotación colisionan frecuentemente contra sus vecinos y se interrumpe el desplazamiento. Por tanto, el movimiento de cada grupo en el interior de la proteína semeja el de una molécula sumergida en un líquido; las colisiones con las moléculas del solvente frenan la molécula de soluto (frenado de arrastre) y generan una trayectoria errática que sigue un "paseo estocástico", el movimiento browniano de difusión.

A lo largo de intervalos prolongados, de varios picosegundos, los rasgos sobresalientes del movimiento se parecen menos al comportamiento de los líquidos y más al de los sólidos. En particular, las fuerzas que mantienen la estructura nativa de la proteína impiden que los átomos se alejen de sus posiciones promedio. Los movimientos propios de un sólido suelen ser de carácter más general que las oscilaciones locales; van desde el desplazamiento simultáneo de unos pocos aminoácidos vecinos hasta deformaciones de la molécula entera. Los efectos de fricción que resultan de las colisiones entre los átomos de la proteína y entre éstos y las moléculas del solvente amortiguan hasta cierto punto ese movimiento. Pese a lo cual, los desplazamientos colectivos que abarcan períodos de varios picosegundos suelen aportar la mayor contribución a las amplitudes globales de las fluctuaciones atómicas de las proteínas. La razón es obvia: ante el denso empaquetamiento del interior de la proteína, un átomo cualquiera no puede alejarse de su posición a menos que también lo hagan sus vecinos. Los movimientos más pronunciados se registran en la superficie de la molécula, donde revisten menor importancia las limitaciones impuestas por el empaquetamiento. En su fluctuación, los átomos del interior llegan a desplazarse hasta medio angstrom, distancia comparable a la del radio atómico; los de la superficie efectúan recorridos de hasta dos angstrom.

La mioglobina, proteína globular que almacena oxígeno en el tejido muscular, constituye un ejemplo de polipéptido en cuya actividad reviste importancia la dinámica molecular. En los cetáceos, por ejemplo, la gran do-





4. CANAL DE ESCAPE de una molécula de oxígeno unida al grupo hemo de una molécula de mioglobina; se abre por la rotación de las cadenas laterales de tres aminoácidos. La imagen superior muestra la estructura, estática, de la mioglobina determinada por cristalografía de rayos X. No existe ninguna ruta que lleve de la superficie de la molécula (*puntos naranja*) a una cavidad del interior de la molécula situada justo sobre el grupo hemo (*azul*). La cavidad se encuentra cerca del lugar donde el hierro del grupo hemo se une a una molécula de oxígeno. En la imagen inferior han rotado las cadenas laterales de tres aminoácidos (*púrpura*), lo que provoca la apertura de un canal de escape. En la simulación no se alteran las posiciones de los restantes átomos del esqueleto ni de las cadenas laterales (*verde*). Si la estructura de la molécula de mioglobina fuera rígida, resultando imposible cualquier rotación de las cadenas laterales, una molécula de oxígeno tardaría miles de millones de años en alcanzar o abandonar la posición de enlace. Las simulaciones sugieren que el oxígeno se enfrenta a una sucesión de barreras, y que avanza entre los átomos de la mioglobina en un ''paseo estocástico'' muy parecido al que sigue una pequeña molécula en un líquido.

tación de mioglobina aporta el oxígeno que precisan en sus prolongadas inmersiones. Fue precisamente la de la mioglobina la primera estructura proteínica que se determinó, por obra de John C. Kendrew, de la Universidad de Cambridge. La proteína que sometió a análisis cristalográfico por rayos X procedía de cachalotes, alimento de los ingleses en la segunda guerra mundial.

En términos generales se conoce bien el modo de acción de la mioglobina. El oxígeno que acarrea la molécula se enlaza de forma reversible a un átomo de hierro situado en el centro de una molécula orgánica compleja y plana: el grupo hemo. Este se encuentra en lo más profundo de la molécula de mioglobina, protegido del agua por la proteína globular que lo rodea, densamente empaquetada. (El agua oxidaría el hierro del grupo hemo, que pasaría del estado ferroso, Fe^{+2}, al férrico, Fe^{+3}, no pudiendo entonces unirse al oxígeno.) El empaquetamiento, sin embargo, es de tal densidad que el oxígeno no puede penetrar hasta el grupo hemo. Si los átomos de la mioglobina se mantuvieran fijos en las posiciones que señala la estructura cristalográfica por rayos X, de poco serviría la mioglobina: el tiempo requerido para que se enlazara un átomo de oxígeno al grupo hemo, o se liberara de nuevo cuando así fuera necesario, superaría en mucho la vida entera del cachalote.

De lo que se ha comentado ya acerca de las fluctuaciones atómicas que se registran en las proteínas cabe intuir por dónde discurre la resolución del problema: las barreras energéticas que impiden la entrada y salida del oxígeno de la rígida estructura proteínica se reducen merced a la suma de movimientos locales. En la exploración de esa posibilidad se han examinado las rutas que con más verosimilitud llevan al grupo hemo, y se han calculado las barreras de esas trayectorias, tanto en la proteína rígida (esto es, la molécula de mioglobina cuya estructura coincide con la que se ha obtenido por cristaligrafía de rayos X) como en el modelo, más realista, cuya estructura puede relajarse a medida que el oxígeno avanza por la matriz proteica. En todas las rutas investigadas, la proteína rígida presentaba barreras energéticas de al menos 100 kilocalorías por mol. El tiempo que tardaría una molécula de oxígeno en entrar o salir a través de una barrera de esa índole sería de varios miles de millones de años.

En la simulación de las posibles fluctuaciones que pudieran abrir paso al oxígeno en su entrada o salida de la mioglobina se tuvieron en consideración los movimientos de las cadenas laterales que generan la barrera de alta energía. En una de las rutas, los obstáculos principales parecían ser las cadenas laterales de tres aminoácidos: histidina *E*7, treonina *E*10 y valina *E*11, según los bautizamos. Simulando la rotación de esas cadenas laterales se calculó la energía necesaria para apartarlas del camino. Basta una energía de 8,5 kilocalorías por mol para dejar expedita la ruta, de modo que la barrera energética para el oxígeno se reduce a unas cinco kilocalorías por mol. El total de energía necesaria pasa, por consiguiente, de 100 kilocalorías por mol a unas 14, cantidad encuadrada en el orden de magnitud que, según sugieren los experimentos, viene a medir la barrera que se opone al enlace del oxígeno.

Por último, para examinar el movimiento del oxígeno se calculó una serie de trayectorias mediante dinámica molecular que podían seguir las moléculas de oxígeno situadas inicialmente en el grupo hemo. Puesto que las distintas trayectorias sólo podían identificarse durante un tiempo limitado, no todas las moléculas de oxígeno lograban escapar; algunas quedaban en un repliegue cercano al grupo hemo y otras iban a parar a distintos puntos del interior de la proteína. Muchas alcanzaron el exterior. En una ruta típica, la molécula de oxígeno debía superar diversos obstáculos, cada uno de los cuales presentaba una barrera energética propia. El oxígeno permanecía largo rato en cada "pozo" de energía, moviéndose de un lado a otro y chocando contra sus paredes. Entonces, al reducirse significativamente la barrera energética por alguna fluctuación de la proteína, o si el oxígeno cobraba suficiente energía por su colisión con los átomos de la proteína, o ambas cosas a la vez (lo más probable), el oxígeno superaba rápidamente la barrera y caía en el pozo siguiente, donde se repetía el proceso.

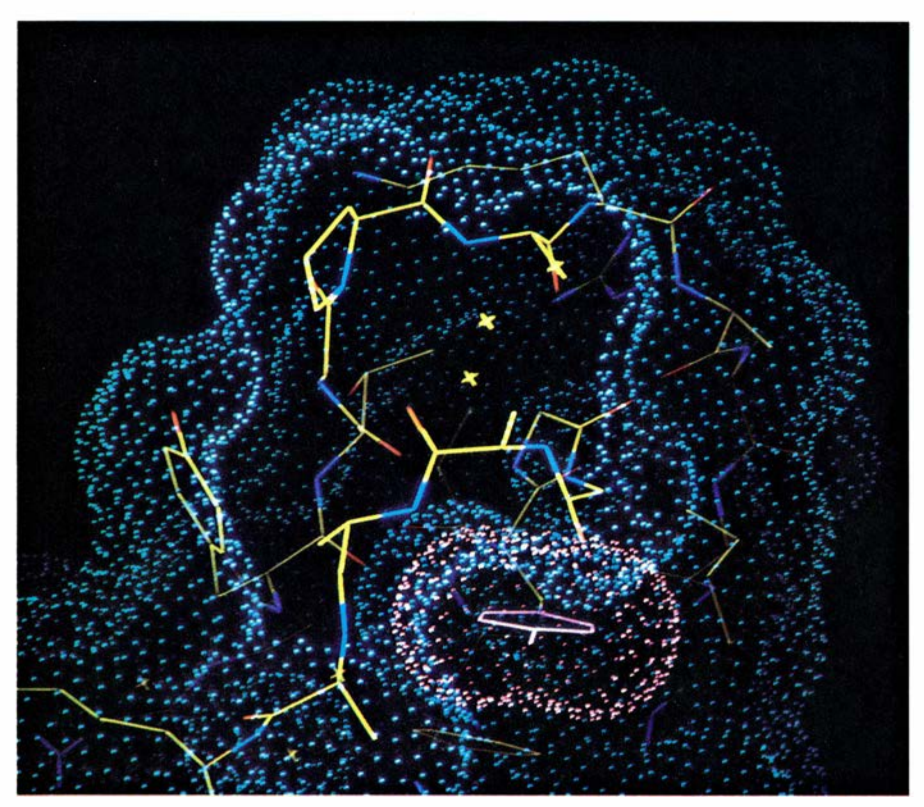
Dada la complejidad de cada vía y el número de rutas posibles, es probable que el movimiento del oxígeno a través de la proteína sea de carácter difusivo. Cuando el oxígeno cruza una barrera no tiene por qué dirigirse necesariamente a la siguiente. Las colisiones pueden imprimir características aleatorias a su dirección, de tal forma que vuelva a cruzar la misma barrera en sentido contrario, equivocando el avance. El comportamiento resultante viene a ser un paseo estocástico parecido al de una molécula de soluto que se difunde por la solución.

El análisis de la mioglobina viene a sugerir que lo que ocurre en esa proteína tiene una validez general. Los ligandos, moléculas que se unen a muchas proteínas, no lograrían entrar ni salir del polipéptido si los átomos de la proteína se encontraran obligados a ocupar sus posiciones promedio. En las enzimas, por tanto, quizá constituya a menudo condición necesaria para que entren y se unan los substratos, y para que salgan los productos de la reacción, el que se registren fluctuaciones de las cadenas laterales o de otras agrupaciones atómicas.

Muchos pasos importantes de la actividad de una proteína abarcan períodos temporales relativamente largos: desde nanosegundos (10^{-9} segundos) hasta milisegundos (10^{-3} segundos), o incluso más. Por ejemplo, la velocidad de la reacción catalizada por algunas enzimas aparece limitada por el tiempo que tardan ciertos grupos atómicos de la enzima en cambiar de una conformación a otra, en la cual pueden intervenir en la catálisis. Cuando se registra, el cambio es rápido; la duración se debe a que los grupos de átomos que participan en la transición se activan sólo muy raramente. Es decir, las fluctuaciones que permiten la transición son acontecimientos infrecuentes.

La simulación convencional de la dinámica molecular tiene poca aplicación en el estudio de los procesos activados; la escasa frecuencia con que se produce la transición conformacional hace muy improbable que se produzca en el intervalo que cubre la simulación. Ahora bien, si se conocen aproximadamente los desplazamientos implicados en la transición, cabe analizar los pormenores del proceso por medio de un método de simulación especializado.

Esa simulación se lleva a cabo en dos pasos. En primer lugar, a partir de una secuencia de simulaciones se calcula la magnitud de la barrera energética que se opone a la transición; tales simulaciones son del tipo habitual, salvo que la molécula sometida a simulación sólo puede moverse en una sucesión limitada de regiones, las que encontrará en la ruta de la transición. Se calculan entonces las trayectorias que parten de configuraciones situadas en la cumbre de la barrera, o en su vecindad. Se elimina así la necesidad de esperar hasta que, por azar, aparezca una molécula



5. ANILLO HEXAGONAL de átomos de carbono que caracteriza la cadena lateral del aminoácido tirosina (*magenta*); aparece aquí en la región interna de la proteína denominada inhibidor de la tripsina pancreática de buey. La imagen muestra los enlaces que constituye el anillo (*líneas magenta*), además de su superficie de van der Waals (*puntos magenta*), que encaja con gran precisión en la estructura estática de la proteína (*puntos azules*). De ser rígida la proteína, el anillo de tirosina no podría dar la vuelta, esto es, rotar 180 grados.

activada. Las trayectorias derivadas de esas simulaciones, junto con el cálculo de la magnitud de la barrera energética, permiten predecir la velocidad a la que ocurrirá la transición.

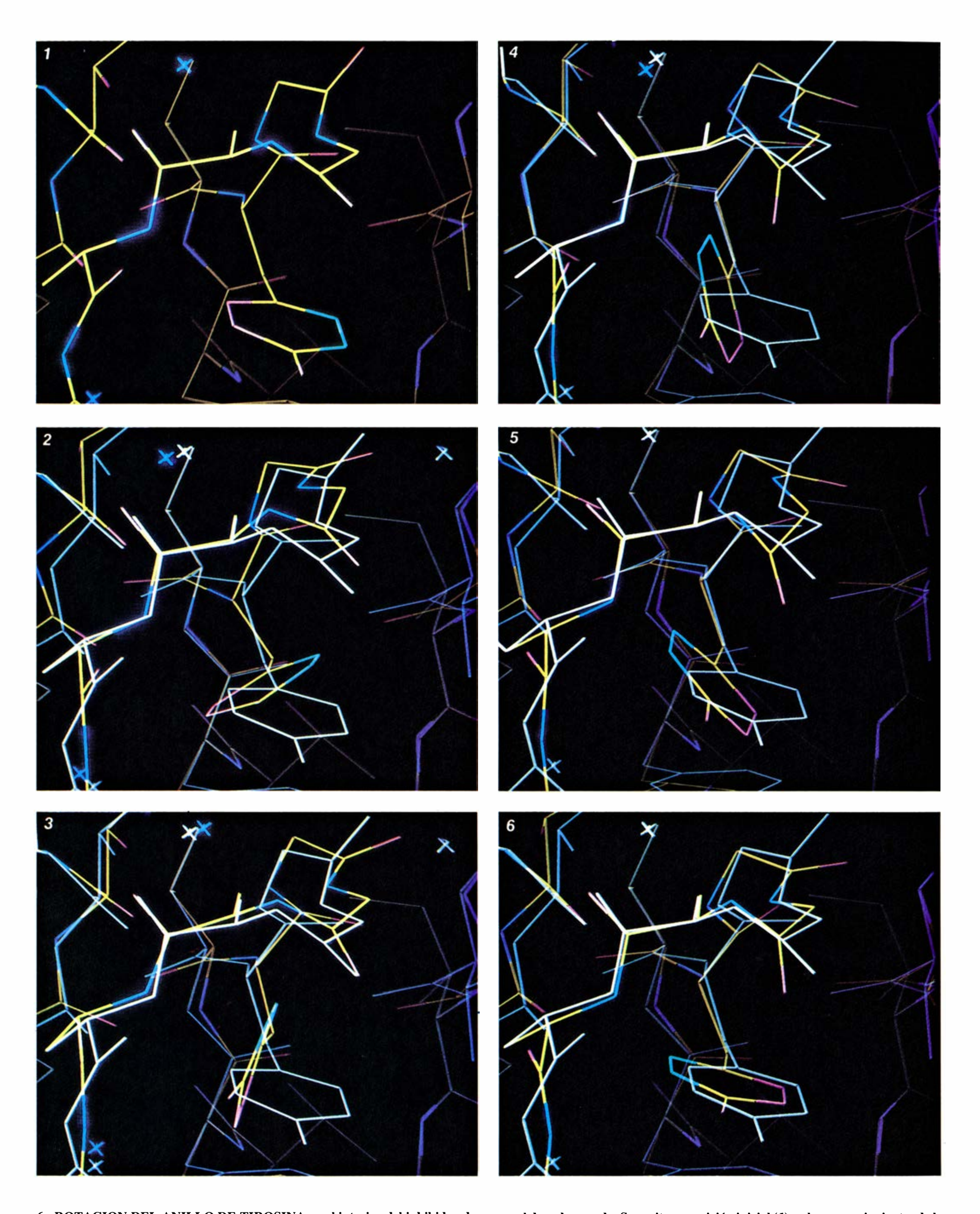
La aplicación de ese método de dinámica activada a algunos procesos sencillos nos ha introducido en la naturaleza de las fluctuaciones locales involucradas en la activación de transiciones de conformación. La transición que describiremos aquí (más por ser representativa de un proceso activado que porque a esa transición le corresponda algún papel biológico importante) es la rotación de un anillo de tirosina situado en las profundidades de la proteína globular denominada inhibidor de la tripsina pancreática bovina. Como su nombre indica, el inhibidor de la tripsina suspende la actividad de esa enzima digestiva uniéndose al centro activo donde la tripsina cataliza la escisión de moléculas de proteínas. Por su parte, la cadena lateral del aminoácido tirosina contiene seis átomos de carbono, que forman un anillo hexagonal plano. Interesan las rotaciones de ese anillo porque se han estudiado experimentalmente por resonancia magnética nuclear. Demuestra la técnica de simulación de la dinámica molecular que la rotación posee carácter colectivo, en el sentido de que la precede el desplazamiento de una sección de la cadena polipeptídica situada sobre una cara del anillo.

El desplazamiento de la cadena contribuye de forma decisiva al mecanismo de rotación; aparta átomos de la ruta del anillo, reduciéndose considerablemente la barrera energética que impide el giro. La energía que precisa la distorsión del esqueleto es mucho menor que la debida al contacto repulsivo entre los átomos del anillo y el esqueleto sin distorsionar. Es más, la fluctuación colectiva facilita el inicio de la rotación: abre un pequeño hueco, hacia el que puede rotar el anillo en respuesta a sus colisiones contra los átomos vecinos restantes. Esa situación favorece la rotación y el anillo acumula la energía cinética necesaria para remontar lo que quede de la barrera energética.

Una sencilla transición conformacional (la rotación de la tirosina) aparece, por tanto, regulada por movimientos colectivos del interior de la proteína. Dada la elevada densidad con que se empaquetan los átomos en las proteínas globulares, es probable que esas transiciones resulten un fenómeno habitual. Sus frecuencias de aparición dependen de ciertas características estructurales de la proteína en una escala relativamente grande, de modo que bien podrían esas transiciones ejercer control de la actividad de la proteína. Por ejemplo, el enlace de una molécula pequeña a un punto de la proteína quizá altere la movilidad de ésta y modifique así la cinética de algún proceso en otra región.

Ejemplo de proteína en cuya actividad enzimática participan movimientos a gran escala es la alcoholdeshidrogenasa hepática, la enzima que oxida el alcohol etílico y colabora en la limitación de los efectos de la ingesta de vino y licores. (Quienes carecen de esa enzima son anormalmente sensibles al alcohol.) Según se cree, el verdadero papel que le corresponde a la enzima es la catalización de otra reacción bioquímica, puesto que no se ven las razones de que la evolución haya desarrollado una enzima de desalcoholización. En todo caso, la alcoholdeshidrogenasa hepática viene a ejemplificar aquellas enzimas o proteínas, así las inmunoglobulinas, que constan de dos o más dominios conectados por filamentos de cadena polipeptídica, que pueden actuar de bisagra. En muchas enzimas, la región catalítica se encuentra entre dos dominios globulares; en ese caso, las deformaciones de la molécula entera por flexión de la bisagra, que acercan o separan los lóbulos, regulan el acceso al centro activo. La velocidad de enlace con el substrato o de liberación de productos quizá dependa, por tanto, de la dinámica de los dominios.

La alcoholdeshidrogenasa es un dímero: su molécula consta de dos subunidades idénticas, dos monómeros. A su vez, cada monómero posee dos dominios lobulares. Carl Bränden, Hans Ecklund y sus colegas, de la Universidad de Uppsala, han revelado, por cristalografía de rayos X, que el monómero presenta dos configuraciones. Una de las estructuras, el denominado apoenzima, es abierta, sus dominios se encuentran separados; la otra, el holoenzima, es cerrada: sus dominios se comprimen junto con un coenzima (molécula auxiliar que interviene en reacciones enzimáticas) unido a uno de los dominios. En este caso, el coenzima es la molécula de transporte de electrones denominada NADH. La estructura abierta permite que el coenzima alcance su región de enlace, mientras que la cerrada proporciona un entorno donde sustrato y coenzima quedan protegidos, pudiendo discurrir la reacción con eficacia (el segundo lóbulo es el dominio catalítico de la enzima). Concluida la reacción, una fluctuación de



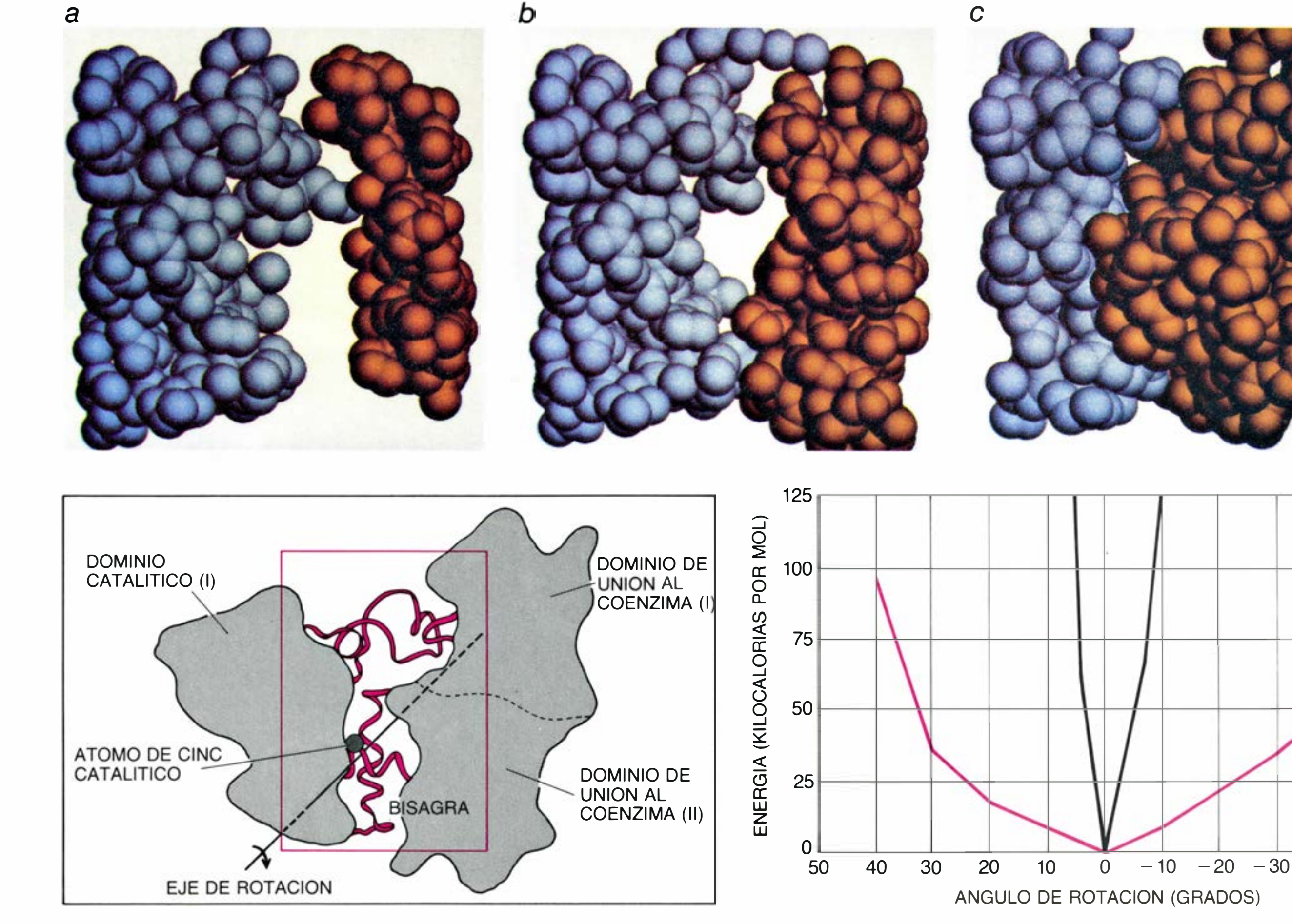
6. ROTACION DEL ANILLO DE TIROSINA en el interior del inhibidor de la tripsina pancreática. Vale de ejemplo de los procesos moleculares en los que las fluctuaciones estocásticas de una molécula de proteína permiten una transición conformacional brusca. El campo visual de esta simulación de la dinámica molecular muestra el anillo y parte del esqueleto polipeptídico que está en contacto con la cara superior del anillo, según recoge la figura 2. Se ha coloreado el anillo para facilitar la observación de su rotación: el borde izquierdo es rojo y, el derecho, verde. Se repite su posición inicial (*1*) en los pasos siguientes de la simulación para que sirva de referencia (*hexágono blanco en 2-6*). Al poco de iniciarse la simulación, el esqueleto se aleja del anillo de tirosina. Tal fluctuación abre un hueco que reduce considerablemente la barrera energética que en un principio se opone al giro; la colisión del anillo con los átomos que persisten en sus inmediaciones tiende a impulsar la rotación. Las imágenes abarcan un intervalo de dos picosegundos, durante el cual se produce una rotación completa.

apertura separa los dominios, escapa el producto de la reacción y la enzima queda lista para atravesar un nuevo ciclo.

En el punto de partida del examen de las fluctuaciones de apertura y cierre se supuso que los dos dominios de un monómero se movían a modo de cuerpos rígidos conectados por una bisagra flexible. Puede someterse a ensayo esa hipótesis superponiendo las estructuras cristalográficas abierta y cerrada (esto es, las configuraciones apo y holo). Al ensayar la coincidencia de los dominios de unión con el coenzima de las formas apo y holo se advierten notables discrepancias en las posiciones de los átomos de los dominios catalíticos. Sin embargo, si se rota rígidamente el dominio catalítico del apoenzima, de modo que la bisagra que une los lóbulos se cierre unos 10 grados, la mayoría de los átomos del dominio catalítico del apoenzima pueden ya superponerse a los átomos equivalentes del correspondiente dominio del holoenzima.

No obstante, resulta una simplificación excesiva suponer que la rotación es rígida. Demuestra un modelo, que incorpora los tipos de interacciones requeridas para calcular las fuerzas que intervienen en una simulación de la dinámica molecular, que la configuración holo obtenida por rotación rígida del dominio catalítico apo posee una energía varios miles de kilocalorías por mol superior a la estructura de la apo. Corresponde esa energía a una inmensa barrera con respecto a la energía térmica disponible: barrera que impediría cualquier rotación. Aparece ese obstáculo por no haberse tenido en cuenta las pequeñas fluctuaciones de ámbito local cuya existencia demuestran las simulaciones de la dinámica molecular. Para que resulte viable la flexión de la bisagra a temperatura ambiente es esencial que se relajen las interacciones atómicas desfavorables generadas por la rotación.

Puede simularse el movimiento relajado de la flexión de la bisagra por medio de un cálculo adiabático del potencial. En ese procedimiento se rota ligeramente (por ejemplo un grado) un dominio del monómero de la alcohol-deshidrogenasa hepática con respecto al otro. Se deja así que los átomos de la región de bisagra y de las zonas de contacto entre los dos dominios experimenten fluctuaciones posicionales que minimicen la energía de la estructura rotada. Se imprime otra rotación de un grado y se repite la minimización. El resultado final, acumulación de la secuencia de rotaciones incrementales y de relajaciones, pondera la energía potencial de la molécula en función del ángulo de rotación de la bisagra. La curva de energía que abarca el intervalo que va desde el apoenzima al holoenzima resulta así bastante plana; de hecho, los movimientos brownianos aleatorios de los lóbulos de la molécula en solución a temperatura ambiente poseerían suficiente energía



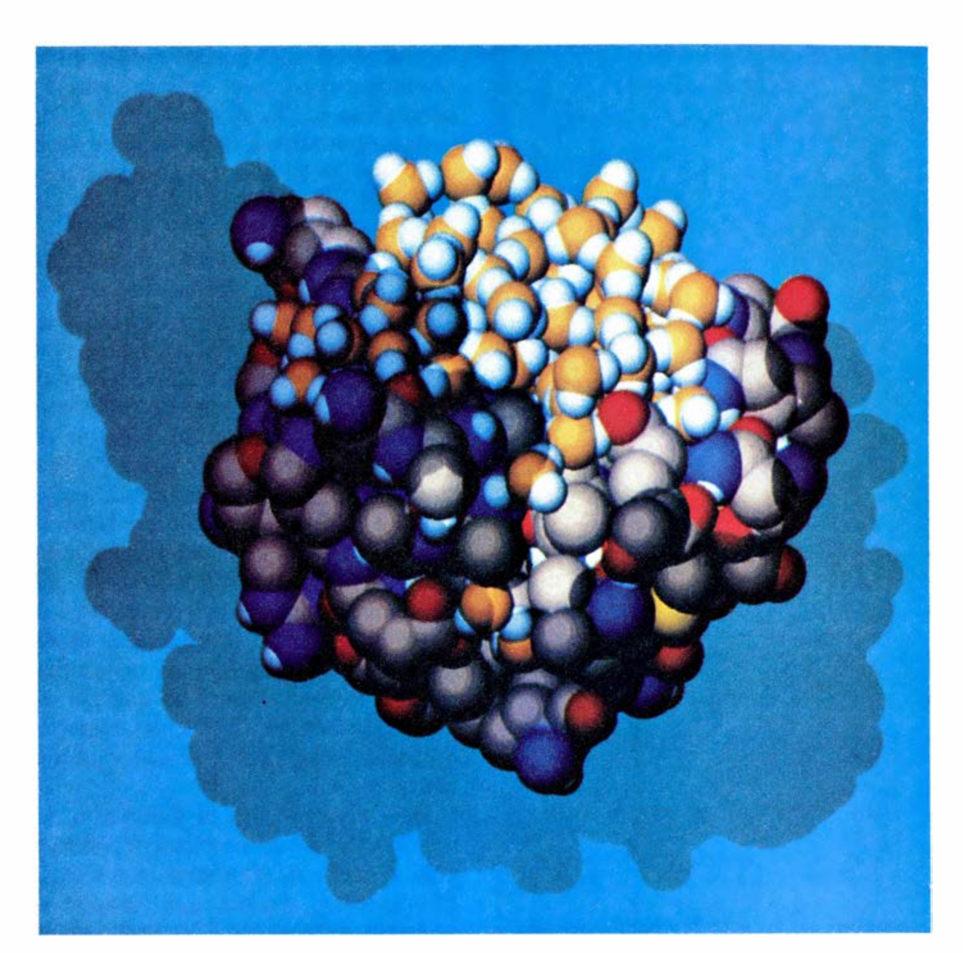
7. MOVIMIENTO DE FLEXION o plegamiento de la bisagra de la enzima alcoholdeshidrogenasa hepática. Acerca los dos dominios de la molécula y crea un ambiente protegido donde transcurre la reacción catalizada por la enzima. Las simulaciones de las porciones de los lóbulos en su proceso de aproximación son obra de Carl Bränden, de la Universidad de Uppsala. Se muestran un ángulo de bisagra de 30 grados (*a*) (bisagra abierta), de 0 grados (*b*) y de −40 grados (*c*) (bisagra cerrada). Cuando el ángulo es nulo la enzima se encuentra en su conformación apo, relativamente abierta, la observada por cristalografía de rayos X. Las simulaciones establecen que si la bisagra fuera rígida (*curva negra de la gráfica inferior*) los movimientos de giro requerirían mucha más energía que la disponible a temperatura ambiente. Por el contrario, si fluctúan (*curva de color*), los movimientos de plegamiento de la bisagra resultan verosímiles.

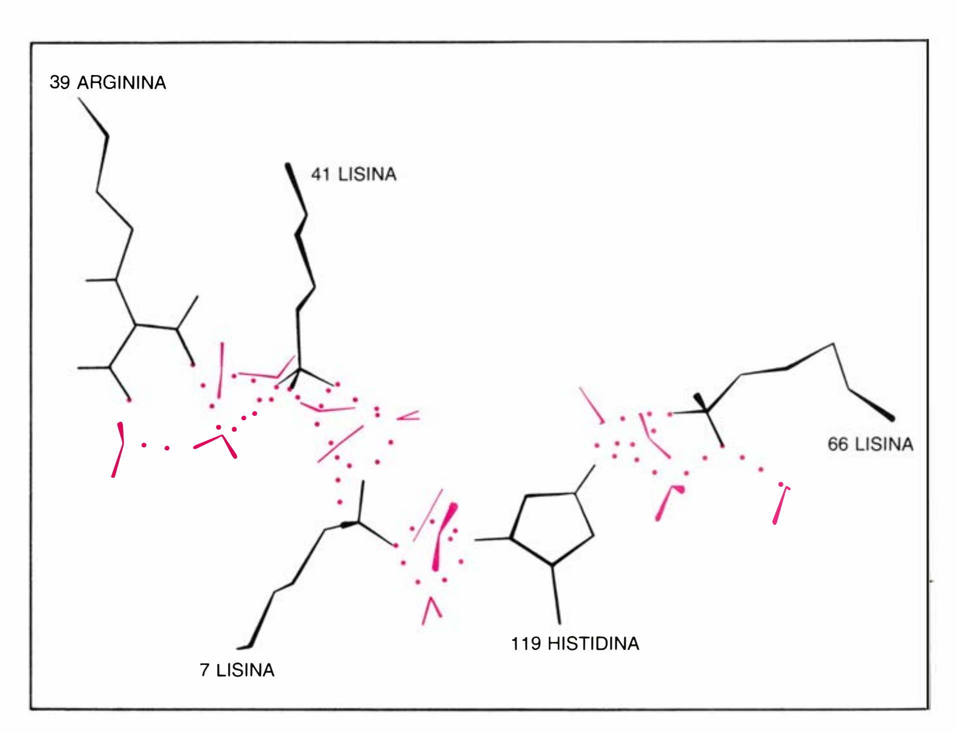
para provocar espontáneamente su cierre y apertura en una escala temporal de nanosegundos.

Los desplazamientos atómicos que dan lugar a la relajación del monómero son, con pocas excepciones, inferiores a los 0,5 angstrom; resultan, por tanto, comparables a los movimientos atómicos normales de la proteína a temperatura ambiente. Tienen lugar, como se dijo antes, en una escala temporal de picosegundos, de modo que son mucho más rápidos que el propio movimiento, en gran escala, de flexión de la bisagra. Lo que justifica la técnica de simulación: todas las alteraciones del ángulo de rotación dan tiempo suficiente para que se introduzcan las relajaciones locales instadas por la minimización de la energía. El análisis de la alcoholdeshidrogenasa hepática (y de otras proteínas con bisagra flexible que se han sometido a estudio teórico) destaca el importante papel que les corresponde a las pequeñas fluctuaciones, de alta frecuencia, en la facilitación de otros movimientos de la proteína, mayores y de carácter más global.

Uno de los objetivos de las simulaciones de la dinámica molecular es proporcionar una descripción detallada y cuantitativa de cómo las enzimas aceleran las velocidades de las reacciones bioquímicas. Se cree que actúan por combinación de varios mecanismos. Se unen a los reactivos para acercarlos entre sí; es más, las enzimas son capaces de situar cargas asociadas a ciertos aminoácidos de tal forma que se facilite la reacción. Sin embargo, no se conoce lo suficiente ninguna reacción enzimática para predecir su velocidad. Los movimientos de los átomos que componen la enzima, los reactivos e incluso las moléculas de agua (el solvente) situadas en el sitio activo o en su vecindad probablemente influyan de forma importante sobre la velocidad. En muchas enzimas el centro activo queda al alcance del solvente durante la reacción, y la propia agua actúa de reactivo; en otras (como es el caso de la alcoholdeshidrogenasa hepática) tiende a excluirse el solvente.

En la catálisis enzimática, como en general en las reacciones químicas, los movimientos atómicos desempeñan un papel decisivo. Sin duda resultará fructífero, por tanto, cualquier enfoque que se fundamente en la dinámica molecular. La simulación por ordenador de una enzima entera y de su solvente resultaría un procedimiento poco eficaz (y caro) para estudiar los movimientos





8. METODO DE SIMULACION CON FRONTERA. Permite que, en vez de simularse la enzima entera encerrada en una "caja" de agua, se simule sólo el centro activo de la enzima, introduciéndose una región fronteriza que confina las moléculas de agua del centro activo y permite el intercambio de energía a través del límite entre el centro y su entorno. En este dibujo de ordenador, obra de Axel Brünger, se muestra el centro activo de la enzima ribonucleasa y las moléculas de agua esenciales que se encuentran en el lugar (*naranja y blanco*). Las sombras que aparecen en el centro activo y en el fondo sugieren el resto de la enzima, que no recibe tratamiento explícito en la simulación. Las moléculas de agua observadas en la simulación (*abajo, líneas de color*) forman una red de puentes de hidrógeno (*líneas punteadas de color*) que interconectan las cadenas laterales del centro activo dotadas de carga. En cada enlace, una molécula de agua y un aminoácido comparten un átomo de hidrógeno. Tales enlaces ayudan a mantener la morfología del centro activo.

locales directamente implicados en una reacción mediada por catalizadores. La simulación de la dinámica molecular permite limitarse a una "zona de reacción" que incluya el centro activo y su vecindad, el substrato y las moléculas de solvente cercanas. No puede ignorarse el resto del sistema, que se sustituye por una región fronteriza, de donde se impide la fuga de las moléculas del solvente por medio de una fuerza de repulsión y en donde se confinan los átomos de la proteína mediante fuerzas que actúan de tal modo que los átomos presentan las fluctuaciones que les correspondería en el caso de que la simulación considerase la proteína entera. Es más, las ecuaciones del movimiento de los átomos del polipéptido y de las moléculas del solvente de la región fronteriza se modifican para incluir términos de acoplamiento de tipo browniano que permitan a esos átomos y moléculas absorber o liberar energía como si les rodeara el resto del sistema.

Ha comenzado ya la aplicación del método de simulación con frontera a gran número de proteínas, incluida la ribonucleasa, enzima que escinde el ácido ribonucleico (ARN) cortando entre dos nucleótidos, las unidades que componen una hebra de ARN. Se ha simulado ya la enzima nativa (es decir, la ribonucleasa sin substratos nucleotídicos unidos a la enzima) y diversos estadios intermedios de la ruta que sigue la enzima en la escisión de la hebra de ARN. Uno de los hallazgos de los cálculos atañe a la estabilidad de las partes de la estructura de la ribonucleasa implicadas en la unión al sustrato e importantes, pues, en la determinación de la especificidad de la enzima.

En este sentido, las simulaciones han demostrado que se crea una red de puentes de hidrógeno entre el sustrato y gran número de aminoácidos de la ribonucleasa; los enlaces de hidrógeno son uniones electrostáticas entre parejas de átomos cargados, total o parcialmente (polares), que comparten un átomo de hidrógeno. En algunas de esas redes participan las cadenas laterales de la asparagina 67, glutamina 69 y asparagina 71 (que constituyen la región de unión de la enzima con el nucleótido denominado adenina) y la treonina 45 y la serina 123 (que forman el del nucleótido citosina). En ausencia de substrato, las posiciones que corresponden a los átomos del sustrato polar en los sitios de unión las ocupan moléculas de agua, que crean una red de puentes de hidrógeno que mimetiza las interacciones del substrato y estabiliza las posiciones de los aminoácidos de la región de enlace. Puesto que las moléculas de agua van saliendo a medida que entra el substrato, la unión y desprendimiento de moléculas de solvente debe constituir un proceso íntimamente ligado al mecanismo de reacción, cuya investigación se verá facilitada con los métodos de la dinámica molecular.

Ciertos aminoácidos (por ejemplo la histidina 119 y la lisina 41) parecen comportarse de forma bastante más flexible. Sus posiciones relativas en la estructura nativa de la ribonucleasa las mantiene principalmente una red de moléculas de agua que actúan de puente, estabilizadas por las cargas netas de las cadenas laterales de esos aminoácidos. Así, esos aminoácidos muestran una relativa libertad de adoptar nuevas posiciones en presencia de un substrato o un producto intermedio de la reacción.

De esa forma probablemente participen con más eficacia en la catálisis. Mucho habrá que trabajar hasta lograr el análisis exhaustivo de la escisión del ARN que cataliza la ribonucleasa, pero los resultados de que se dispone ya demuestran la importancia de completar los datos obtenidos por cristalografía de rayos X con simulaciones de dinámica molecular.

Los métodos y los resultados que se han esbozado aquí permiten predecir el futuro que aguarda a las simulaciones de dinámica molecular. Listo para su estudio hay un amplio abanico de problemas biológicos relacionados con las proteínas (por no citar los ácidos nucleicos ni los lípidos de la membrana celular). En los próximos años los investigadores deberían aprender a calcular la velocidad de las reacciones enzimáticas, la unión entre moléculas pequeñas y grandes y otros muchos procesos biológicos de importancia; se comprenderá con mucho más detalle el papel que desempeñan la flexibilidad y las fluctuaciones en la actuación de las macromoléculas; debería poderse determinar cómo generan ciertos patrones de fluctuaciones en proteínas determinadas características del solvente, o determinadas secuencias de aminoácidos. A medida que aumente la capacidad de predicción, el estudio de la dinámica de las proteínas hallará aplicación en problemas prácticos de ingeniería genética o en la modificación de enzimas para mejorar su rendimiento en procesos industriales.

Aplicaciones del efecto Hall cuántico

Esta variante de un fenómeno clásico posibilita, incluso en muestras irregulares, la medición de constantes fundamentales con una exactitud que rivaliza con la de las calibraciones más precisas

Bertrand I. Halperin

Pocas áreas de la ciencia moderna se han estudiado con la intensidad de la física de los materiales semiconductores. Son éstos componentes esenciales de los aparatos electrónicos, emblema de nuestra era. Con el refinamiento de estos aparatos, los investigadores han ido dominando el crecimiento de cristales semiconductores rayanos en la perfección y han aprendido a modificar las propiedades electrónicas de tales cristales mediante la adición de unas pocas partes por millón de átomos extraños. El profundo conocimiento de las propiedades de los cristales modificados ha posibilitado su incorporación en elementos del circuito cuyo grosor se mide en decenas de capas atómicas y cuya área es tan pequeña que en una pastilla ("chip") de un centímetro cuadrado se alojan hasta un millón de elementos.

Podría, por tanto, parecer improbable que en los años ochenta los físicos de semiconductores llegaran a verse sorprendidos ante el descubrimiento experimental de un efecto nuevo y revolucionario, tan poco previsible que habrían de transcurrir muchos meses antes de que los expertos en el campo pudiesen elaborar una explicación satisfactoria. De hecho, se han producido dos sorpresas de este tipo.

En 1980, Klaus von Klitzing, que a la sazón trabajaba en el laboratorio de campos magnéticos intensos del Instituto Max Planck en Grenoble, hizo un descubrimiento que iba a reportarle el premio Nobel de física de 1985: observó un fenómeno ahora conocido con el nombre de efecto Hall cuántico entero [véase "El efecto Hall cuántico", de Klaus von Klitzing: INVESTIGACIÓN Y CIENCIA, mayo de 1986]. Dos años más tarde, en un experimento realizado en el laboratorio nacional de electroimanes de Cambridge, en Massachussetts, Daniel C. Tsui, Horst L. Störmer y Arthur C. Gossard, de los laboratorios AT & T Bell, observaron otro fenómeno notable, ahora conocido como efecto Hall cuántico fraccionario. Como sus nombres indican, estos efectos tienen mucho en común, pero requieren explicaciones distintas.

El efecto Hall cuántico, entero y fraccionario, constituye, en realidad, un caso especial de un fenómeno más general conocido por efecto Hall. Para observar este último, se sitúa una muestra laminar de un material conductor en un campo magnético perpendicular a la superficie de la muestra. Mediante electrodos colocados a cada extremo de la muestra, se hace luego pasar una débil corriente eléctrica de un extremo a otro de la misma. (Una corriente eléctrica no es más que el movimiento de partículas cargadas; en el caso de la mayoría de los sólidos, las partículas portadoras de corriente son los electrones.)

Se realizan entonces dos mediciones. Para acometer la primera, a un voltímetro se conectan dos contactos cercanos a ambos extremos de la muestra; el voltímetro mide la caída de potencial que se produce de un extremo de la muestra al otro, es decir, en la dirección paralela al flujo de la corriente. (Se trata, a grandes rasgos, de una medida de la diferencia entre la energía de los electrones que entran en la muestra y la energía de los que la han atravesado.) La otra medición se efectúa con un voltímetro conectado a electrodos situados en los bordes opuestos de la muestra, de suerte que una línea trazada entre los electrodos sería perpendicular a la dirección de la corriente. Este voltímetro mide la diferencia de potencial a través de la muestra, en dirección perpendicular a la corriente.

Si no hubiese campo magnético alguno, la diferencia de potencial perpendicular a la corriente sería nula. Tal como el físico norteamericano Edwin H. Hall observó en 1879, cuando hay un campo magnético perpendicular a la muestra, se produce una diferencia de potencial que, en la mayoría de los casos, es directamente proporcional a la intensidad del campo magnético. La caída de potencial paralela a la corriente es aproximadamente la misma, exista o no campo magnético.

La caída de potencial paralela a la corriente y la caída de potencial perpendicular a la misma son directamente proporcionales a la intensidad de la corriente que fluye por la muestra. Conviene, pues, dividir cada uno de estos potenciales por la intensidad de la corriente que fluye, y así obtener valores que son independientes del flujo de corriente. El cociente entre la caída de potencial a lo largo de la muestra y la intensidad de la corriente constituye la resistencia ordinaria de la muestra. El cociente entre la caída de potencial a través de la muestra y el flujo de corriente a lo largo de ella se denomina resistencia de Hall.

En su experimento de 1980, von Klitzing estaba estudiando el efecto Hall en ciertos dispositivos en los que los electrones que están libres para transportar corriente se hallan confinados en una lámina delgada. Encontró que, cuando tales dispositivos se enfriaban aproximadamente a un grado por encima del cero absoluto y se colocaban en campos magnéticos extremadamente intensos, el comportamiento de la resistencia ordinaria y el de la resistencia de Hall divergían radicalmente del comportamiento descrito por Hall. Había descubierto el efecto Hall cuántico entero.

Las diferencias más sorprendentes entre el efecto Hall cuántico entero y

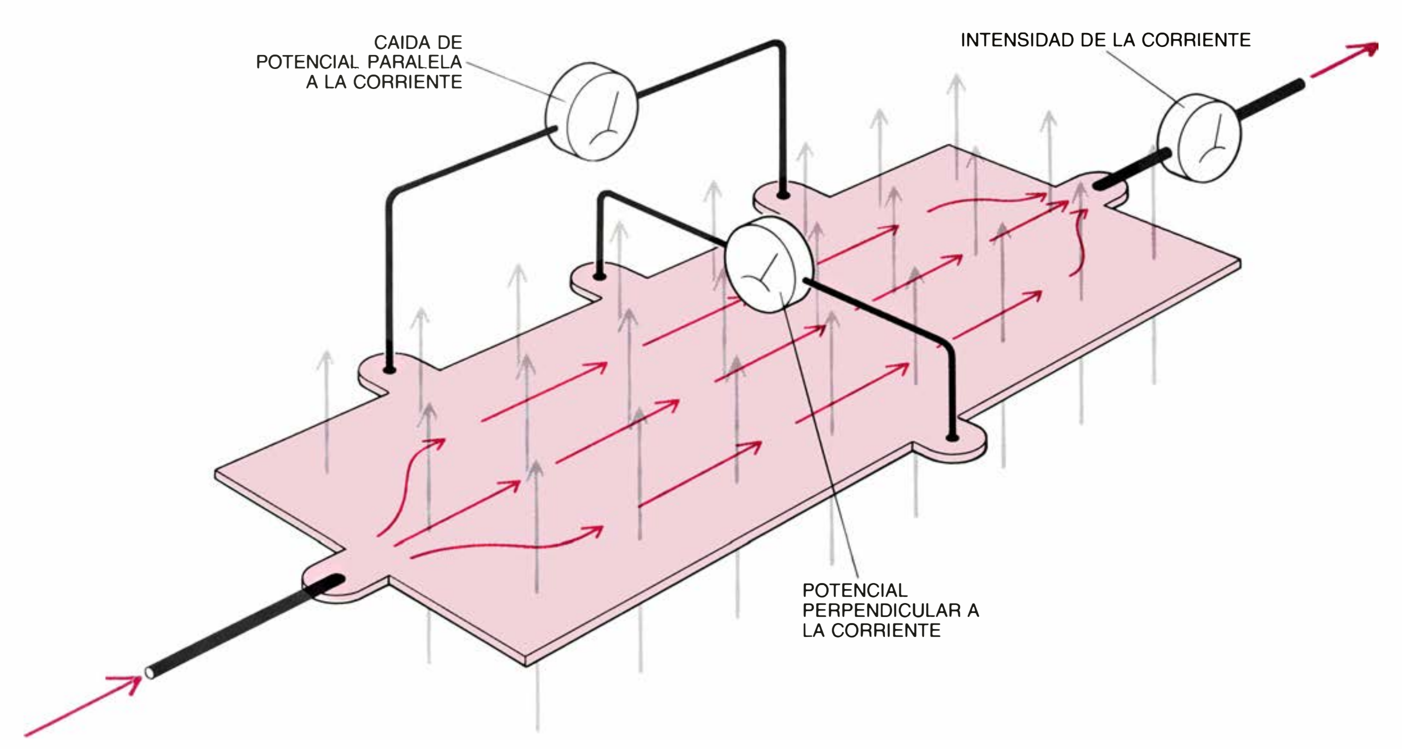
el efecto Hall normal estriba en la resistencia de Hall; en vez de crecer continua y linealmente cuando la intensidad del campo magnético aumenta, el primero exhibe una serie de mesetas ("plateaus"). Es decir, existen intervalos en los que se observa que la resistencia de Hall no varía en absoluto al hacerlo la intensidad del campo magnético. (Entre tales mesetas, la resistencia de Hall crece suavemente al incrementarse el campo magnético.) Además, para los mismos intervalos de la intensidad del campo magnético en que la resistencia de Hall exhibe las mesetas, la caída de potencial paralela a la corriente desaparece por lo que se ve. Es decir, en la muestra no hay resistencia eléctrica alguna y la corriente fluye de un extremo a otro sin la más mínima disipación de energía.

La resistencia eléctrica nula y las mesetas de la resistencia de Hall son fenómenos que merecen atención por sí mismos. El efecto resulta más sorprendente si se examinan los valores para los cuales aparecen las mesetas en la resistencia de Hall. En cada meseta ("plateau") el valor de la resistencia de Hall satisface una condición de notable sencillez: el recíproco de la resistencia de Hall es igual a un entero multiplicado por el cuadrado de la carga del electrón y dividido por la constante de Planck, la constante fundamental de la mecánica cuántica. Cada meseta ("plateau") se caracteriza por un entero distinto.

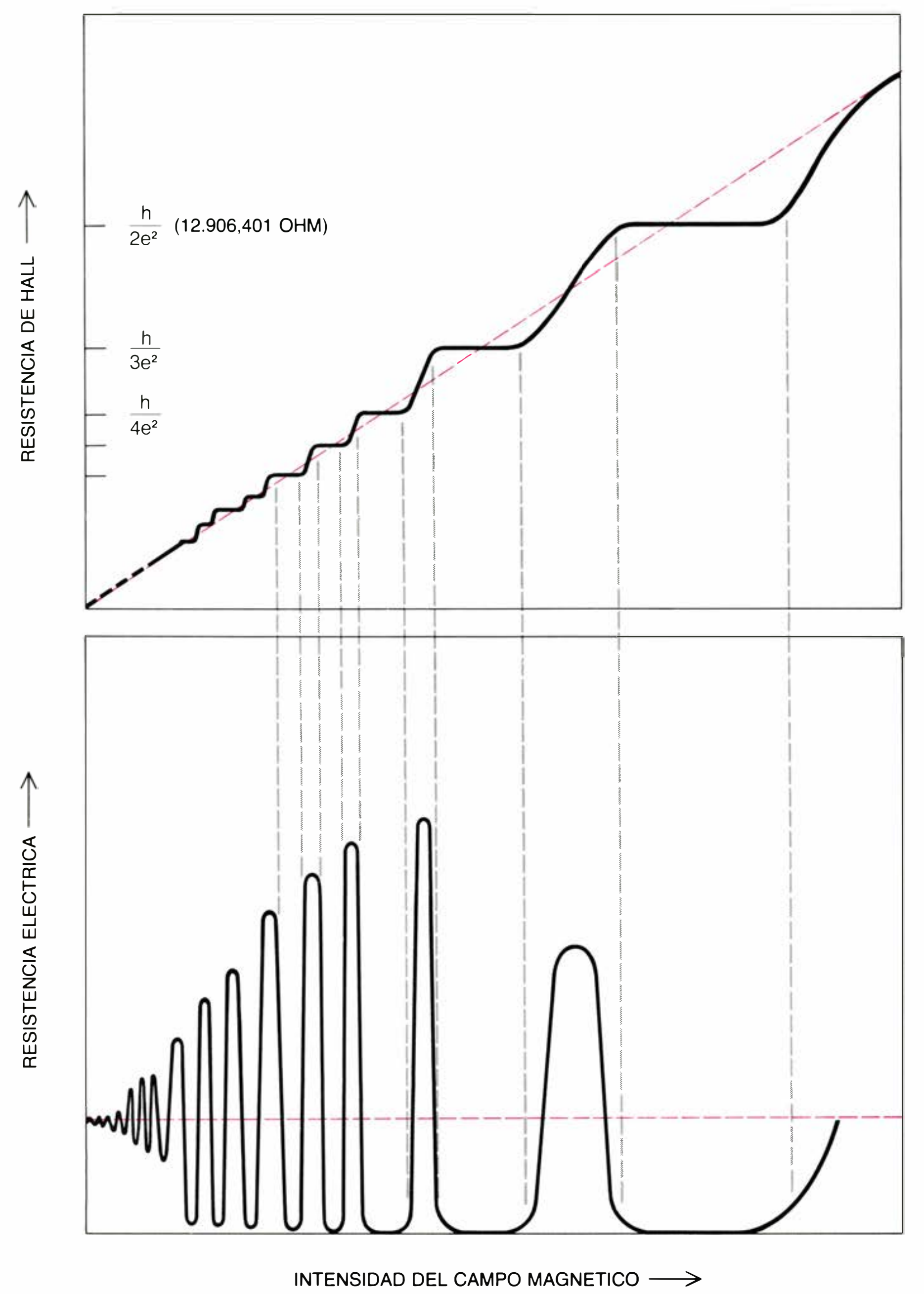
El efecto Hall cuántico fraccionario, descubierto por Tsui, Störmer y Gossard, tiene lugar únicamente en dispositivos exentos, en alto grado, de defectos que interfieran con el movimiento de los electrones. En el efecto fraccionario, existen mesetas adicionales para los que el recíproco de la resistencia de Hall es igual a una fracción simple (no a un entero) multiplicada por el cuadrado de la carga del electrón y dividida por la constante de Planck. En los experimentos originales, las fracciones observadas fueron 1/3 y 2/3. En experimentos más recientes se han descubierto en niveles que corresponden a otras fracciones; 4/3, 5/3, 2/5, 3/5 y 3/7; no se ha hallado, sin embargo, ninguna meseta en niveles correspondientes a fracciones con denominadores pares.

Lo que más sorprende, en lo que concierne al efecto Hall cuántico, es su grado de precisión. Los valores medidos de la resistencia de Hall en varias mesetas satisfacen la relación simple que se ha descrito con una exactitud de alrededor de una parte en seis millones, que es la precisión con la que se conocen los propios valores de las constantes fundamentales. El hecho de que la resistencia de Hall en el efecto entero esté cuantizada (en otras palabras, que los valores individuales de la resistencia de Hall sean iguales a una misma cantidad multiplicada por diferentes enteros) se ha establecido determinando el cociente de las resistencias de Hall en diferentes mesetas y averiguando que son cocientes de números enteros. La exactitud de los experimentos que han determinado tales cocientes llegan hasta la razón de uno a 30 millones.

La precisión de la resistencia de Hall cuántica reviste especial interés si recordamos que, en condiciones normales, la resistencia de Hall de una muestra, al igual que la resistencia eléctrica ordinaria, se manifiesta muy sensible a la geometría de la muestra, su proce-



1. MECANISMO para medir el efecto Hall. Consta de una lámina plana de material conductor y dispositivos para medir potenciales y corrientes. Una determinada cantidad de corriente (*flechas horizontales*), cuya intensidad se ha medido, circula de un extremo a otro de la muestra mientras se aplica un campo magnético (*flechas verticales*) perpendicularmente a la muestra. Se miden dos potenciales: la caída de potencial de uno a otro extremo de la muestra (que constituye, a grandes rasgos, la diferencia de energía entre los electrones que entran en la lámina y los que salen de la misma) y el potencial perpendicular a la corriente, denominado potencial de Hall. En el efecto Hall clásico, la resistencia de Hall (que es igual al potencial de Hall dividido por la intensidad de la corriente que fluye) es directamente proporcional a la intensidad del campo magnético: crece de forma continua con la intensidad del campo. En el efecto Hall cuántico, que se observa cuando los electrones que transportan la corriente están confinados en una delgada lámina de un dispositivo semiconductor mantenido a muy bajas temperaturas y en campos magnéticos extremadamente intensos, la resistencia de Hall describe mesetas ("plateaus"): existen intervalos de la intensidad del campo magnético para los que la resistencia de Hall permanece constante al variar el campo magnético. Además, en estos intervalos, la resistencia eléctrica de la muestra (la caída de potencial paralela a la corriente dividida por la intensidad de la corriente) desaparece por completo, lo que indica que la corriente fluye a través de la muestra sin que se produzca la más mínima disipación de energía. Los valores de la resistencia de Hall en las diferentes mesetas son siempre exactamente los mismos, con independencia (dentro de límites muy amplios) de cuál sea la geometría e incluso la composición de la muestra.



2. EFECTO HALL CUANTICO. Se manifiesta a través de mesetas en la resistencia de Hall de una muestra (*arriba*), que coinciden con la desaparición de la resistencia eléctrica de la misma (*abajo*). En las mesetas, la resistencia de Hall no cambia aunque varíe la intensidad del campo magnético. En los experimentos clásicos (*color*), sin embargo, la resistencia de Hall aumenta linealmente al incrementarse la intensidad del campo magnético, mientras que la resistencia ordinaria permanece constante. En cada meseta el valor de la resistencia de Hall es exactamente igual a la constante de Planck (la fundamental de la mecánica cuántica, *h*) dividida por un múltiplo entero de la carga del electrón al cuadrado. De este modo, el efecto Hall cuántico permite mediciones precisas de las constantes fundamentales y calibraciones sutiles de los instrumentos.

sado, la cantidad de impurezas que contiene y su temperatura en el momento del experimento, entre otros pormenores. El efecto Hall cuántico, por el contrario, es el mismo para muestras de distintas geometrías a diferentes procesados e incluso para las que se han fabricado a partir de materiales diversos. Apenas si se puede hallar precedente alguno en lo que respecta a propiedades de un sólido que sean reproducibles de una muestra a otra con el nivel de precisión alcanzado por el efecto Hall cuántico.

Ambos tipos de efecto Hall cuántico, entero y fraccionario, se detectan en muestras semiconductoras donde los electrones que están libres para transportar la corriente se hallan confinados en una delgada lámina del semiconductor. Hay dos dispositivos ordinarios que confinan a los electrones de esta manera. En el dispositivo original estudiado por von Klitzing, denominado transistor de efecto de campo, una delgada lámina de material aislante separa, de un substrato de silicio, un contacto metálico. La tensión aplicada al contacto metálico puede atraer a los electrones hacia una delgada capa, debajo mismo de la superficie del silicio. (Las muestras de alta calidad empleadas por von Klitzing fueron construidas por Gerhardt Dorda, del laboratorio de investigación de la empresa Siemens en Munich, y por Michael Pepper, de la Universidad de Cambridge.)

El dispositivo en el que se observó, por primera vez, el efecto Hall cuántico fraccionario, denominado heterounión, es una yuxtaposición de cristales constituida por dos materiales semiconductores distintos. En una heterounión, los electrones de uno de los semiconductores son atraídos hacia localizaciones del otro semiconductor más favorables desde el punto de vista energético. Sin embargo, la carga positiva así creada en el semiconductor "donante" engendra una fuerza que retrotrae a los electrones; éstos quedan atrapados en una fina lámina, en la cara que separa ambos cristales.

En los transistores de efectos de campo y en las heterouniones, los electrones que transportan la corriente se mueven libres hacia cualquier lugar, siempre dentro del plano donde se hallan confinados. Sin embargo, el movimiento perpendicular a este plano resulta punto menos que imposible y, por consiguiente, a efectos prácticos, la lámina de electrones puede considerarse bidimensional. En los experimentos sobre efecto Hall cuántico, esta lámina bidimensional de electrones desempeña la misma función que la placa delgada de conductor en los experimentos ordinarios del efecto Hall.

Aunque las propiedades de las láminas bidimensionales de electrones se habían estudiado exhaustivamente, el descubrimiento del efecto Hall cuántico entero constituyó una sorpresa para todos. Cierto es que se habían predicho aspectos del efecto, como resultado incidental, en un artículo relativo a un tema afín, escrito por Tsuneya Ando, Yukio Matsumoto y Yasutada Uemura, de la Universidad de Tokio, pero el aspecto más importante, la extrema precisión del efecto, no se esperaba porque la predicción estaba basada en un tipo de análisis que, en general, es sólo aproximadamente correcto para los electrones que se hallan en materiales reales. Para el efecto Hall cuántico fraccionario no había, en

el momento de su descubrimiento, ninguna explicación en la bibliografía.

A pesar de contar hoy con explicaciones convincentes sobre la exactitud del efecto Hall cuántico, no es posible presentar aquí una justificación técnica del efecto Hall cuántico fraccionario, ni tan siquiera del entero. Tal explicación debería estar necesariamente escrita en el lenguaje altamente matemático de la mecánica cuántica, base de todas las descripciones modernas de los fenómenos a escala atómica. Pero sí cabe ofrecer una explicación parcial que incorpora muchas de las ideas esenciales y que, cuando menos, hará plausibles los resultados experimentales.

Los conceptos de campo eléctrico y campo magnético, así como otros clásicos y fundamentales, son esenciales en cualquier exposición del efecto Hall. El campo eléctrico ejerce una fuerza sobre los objetos dotados de carga, de manera análoga a como el campo gravitatorio la ejerce sobre los objetos dotados de masa. Cuando un bloque de mármol cae rodando por una ladera tiende a hacerlo en la dirección en la que la pendiente de la ladera sea máxima. De forma parecida, una partícula cargada en un campo eléctrico siente, en general, una fuerza en la dirección de las líneas de campo.

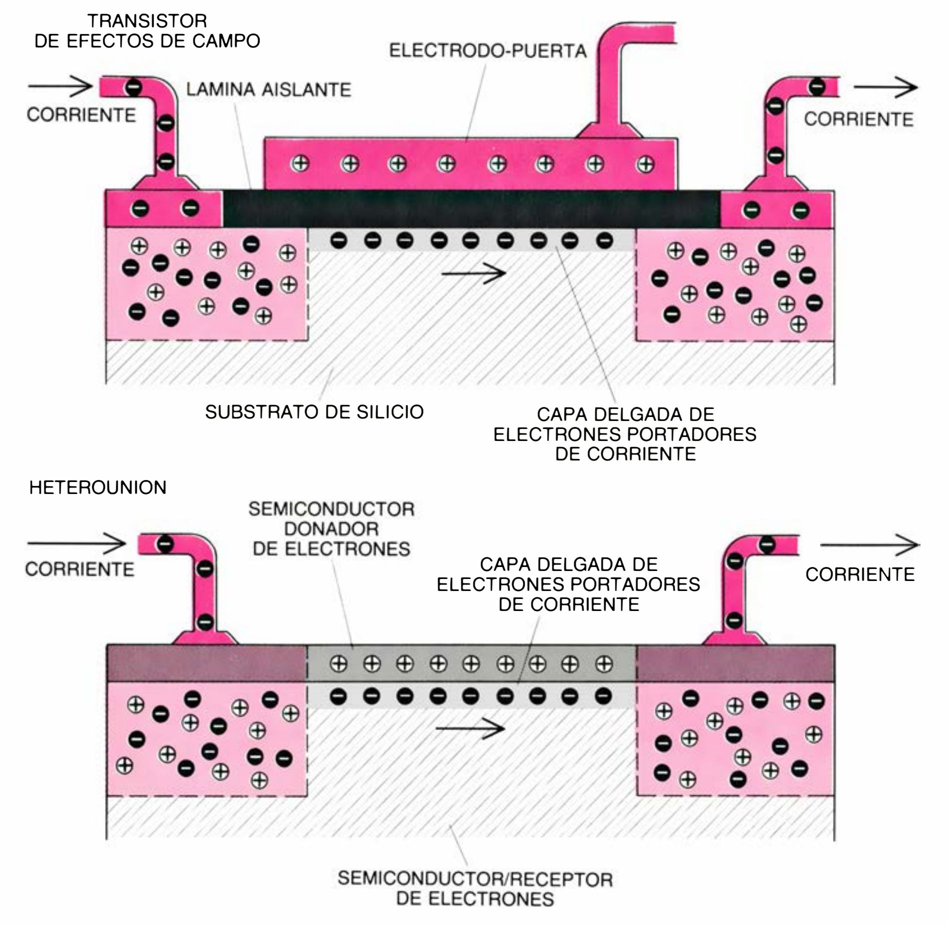
Sigamos con las semejanzas. El campo gravitatorio puede crearse con una concentración de objetos masivos; el eléctrico, con una concentración de partículas cargadas. Así como un objeto que se encuentra en la cima de una colina tiene mayor "energía potencial gravitatoria" que otro alojado en el valle, así también una partícula cargada puede tener mayor "energía potencial eléctrica" que una partícula similar situada en cualquier otro lugar de un determinado campo eléctrico. La diferencia entre la energía potencial eléctrica de partículas cargadas similares, situadas en dos puntos distintos del espacio, recibe el nombre de diferencia de potencial entre estos puntos.

Pero no todo son parecidos. El campo magnético ejerce su fuerza únicamente sobre partículas cargadas en movimiento, no así el campo eléctrico. La magnitud de la fuerza ejercida sobre una partícula es directamente proporcional a la velocidad de la partícula y a la intensidad del campo magnético. La dirección de la fuerza es perpendicular a la dirección de movimiento de la partícula y a la dirección del campo magnético. En consecuencia, una partícula cargada que se mueva en un campo magnético describirá una circunferencia: en cada instante, el campo magnético ejercerá una fuerza perpendicular al movimiento de la partícula y, por tanto, en cada instante la trayectoria de la partícula se curvará. El radio de la órbita circular será inversamente proporcional a la intensidad del campo magnético, de manera que campos magnéticos más intensos mantienen a las partículas en órbitas menores.

Supongamos que una partícula cargada en movimiento se coloca simultáneamente en un campo eléctrico y en un campo magnético. En este caso, la partícula sigue una trayectoria cicloidal: se mueve sobre una circunferencia en torno al denominado centro guía, que, a su vez, se desenvuelve sobre una línea recta [*véase la figura 4*]. Resulta que el centro guía del movimiento de la partícula describe una línea perpendicular al campo eléctrico y al campo magnético. En otras palabras, no se mueve ni hacia arriba ni hacia abajo sobre la "ladera" de la energía potencial eléctrica. Antes bien, en promedio, tiende a moverse horizontalmente a lo largo de la ladera, en dirección perpendicular a la de su máxima pendiente.

La velocidad de movimiento del centro guía, llamada velocidad de deriva, es inversamente proporcional a la intensidad del campo magnético –cuando el campo magnético es más intenso, el movimiento de la partícula es más circular y menos lineal– y directamente proporcional a la intensidad del campo eléctrico; así, el centro guía se mueve más deprisa en las regiones donde la "pendiente" es más pronunciada, aun cuando lo haga a lo largo de ella, y no hacia abajo.

Para entender el efecto Hall clásico nos basta con estos principios básicos. Supongamos que una corriente de electrones fluye de izquierda a derecha en una lámina de metal finita que yace sobre un plano paralelo a esta página. Supongamos ahora que se aplica un campo magnético perpendicular-



3. SISTEMAS SEMICONDUCTORES en los que se observa el efecto Hall cuántico. Mantienen los electrones portadores de corriente dentro de una fina lámina de cristal semiconductor. En un transistor de efectos de campo (*arriba*), los electrones son atraídos hacia la superficie de un cristal de silicio por una carga positiva aplicada a un contacto metálico: el electrodo-puerta. El material aislante impide que los electrones entren en el electrodo; y así permanecen en una capa delgada adyacente a la superficie del cristal. En una heterounión (*abajo*), formada por yuxtaposición de semiconductores con propiedades distintas, los electrones de un semiconductor son atraídos hacia estados de energía más baja del otro. Los iones positivos que dejan tras de sí atraen a los electrones, impeliéndoles a regresar, por lo que quedan atrapados en una delgada capa.

mente a la lámina metálica. Al principio, los electrones serán empujados en una dirección perpendicular a la dirección original y a la del campo magnético: se moverán hacia el borde superior o inferior de la página. Si los electrones sólo pueden abandonar el metal a través de los contactos situados en los extremos izquierdo y derecho, se acumularán en un borde de la lámina (el superior o el inferior). Esta acumulación de partículas cargadas produce un campo eléctrico que arranca de un borde del metal y procede en dirección hacia el otro. Vale decir, uno de los bordes tendrá mayor voltaje que el otro.

El campo eléctrico, en combinación con el magnético, induce que los electrones se muevan en trayectorias cicloidales en una dirección perpendicular a ambos campos: de izquierda a derecha. La diferencia de potencial entre los dos bordes del metal es, por consiguiente, perpendicular a la dirección en la que circula la corriente.

Si para una intensidad de corriente fijada de antemano (es decir, en el caso de que los electrones sigan fluyendo de izquierda a derecha con la misma intensidad), se aumenta la intensidad del campo magnético, entonces la intensidad del campo eléctrico –y con ella la diferencia de potencial entre ambos bordes del metal– debe también incrementarse. La razón estriba en que la "velocidad de deriva" con la que se mueve de izquierda a derecha el centro guía correspondiente a cada trayectoria cicloidal de un electrón es inversamente proporcional a la intensidad del campo magnético y directamente proporcional a la intensidad del campo eléctrico. En consecuencia, cuando aumenta la intensidad del campo magnético, un número mayor de electrones se desvía hacia un borde del metal, creándose un campo eléctrico más intenso y una mayor diferencia de potencial entre ambos bordes. La diferencia de potencial, que aumenta linealmente con la intensidad del campo magnético, constituye el efecto Hall clásico.

Este tipo de análisis clásico no permite explicar el efecto Hall cuántico, que es, por su propia naturaleza, un fenómeno mecánico-cuántico. Es posible, sin embargo, comprender algunos aspectos del efecto Hall cuántico mediante lo que se denomina análisis semiclásico, mezcla de ideas clásicas y mecánico-cuánticas que elude la complejidad matemática de una descripción mecánico-cuántica completa. Aunque un análisis semiclásico no sea del todo fiable, resulta, sin embargo, eficaz como guía que sugiere el esquema general de la respuesta mecánico-cuántica; a menudo, los propios expertos en este dominio se apoyan en él con tal propósito.

En el caso clásico, un electrón sometido a un campo magnético uniforme describe una órbita circular cerrada.

En mecánica cuántica, la energía de una partícula en una órbita cerrada de este tipo puede adquirir únicamente ciertos valores discretos, del mismo modo que un electrón en un átomo solamente puede ocupar ciertos niveles energéticos. Por niveles de Landau hay que entender los niveles de energía del movimiento bidimensinal de una partícula cargada, un electrón, por ejemplo, en un campo magnético uniforme. No hay ningún conjunto discreto de niveles de energía para sistemas "gruesos", tridimensionales, en un campo magnético uniforme, porque la energía del movimiento de la partícula paralelo al campo magnético, a diferencia de lo que sucede con el movimiento en el plano perpendicular al campo, no está cuantizada. Resulta, pues, decisivo, por lo que hace al presente análisis, que el sistema físico en consideración consista en una capa delgada, "bidimensional", de electrones.

En el caso clásico, abundan las regiones del plano donde las órbitas circulares idénticas de los electrones pueden quedar centradas sin que se solapen unas con otras. De manera similar, en el caso semiclásico existe cierto número de estados con el mismo nivel energético que son independientes unos de otros. (Para cada estado hay una "función de onda" distinta, es decir, una descripción matemática distinta de la posición probable del electrón.) En concreto, el número de estados independientes por unidad de área es directamente proporcional a la intensidad del campo magnético. Esto es, cuando menos, plausible, dado que en el caso clásico un campo magnético más intenso confina a los electrones en órbitas menores y, en consecuencia, en un área determinada pueden caber más

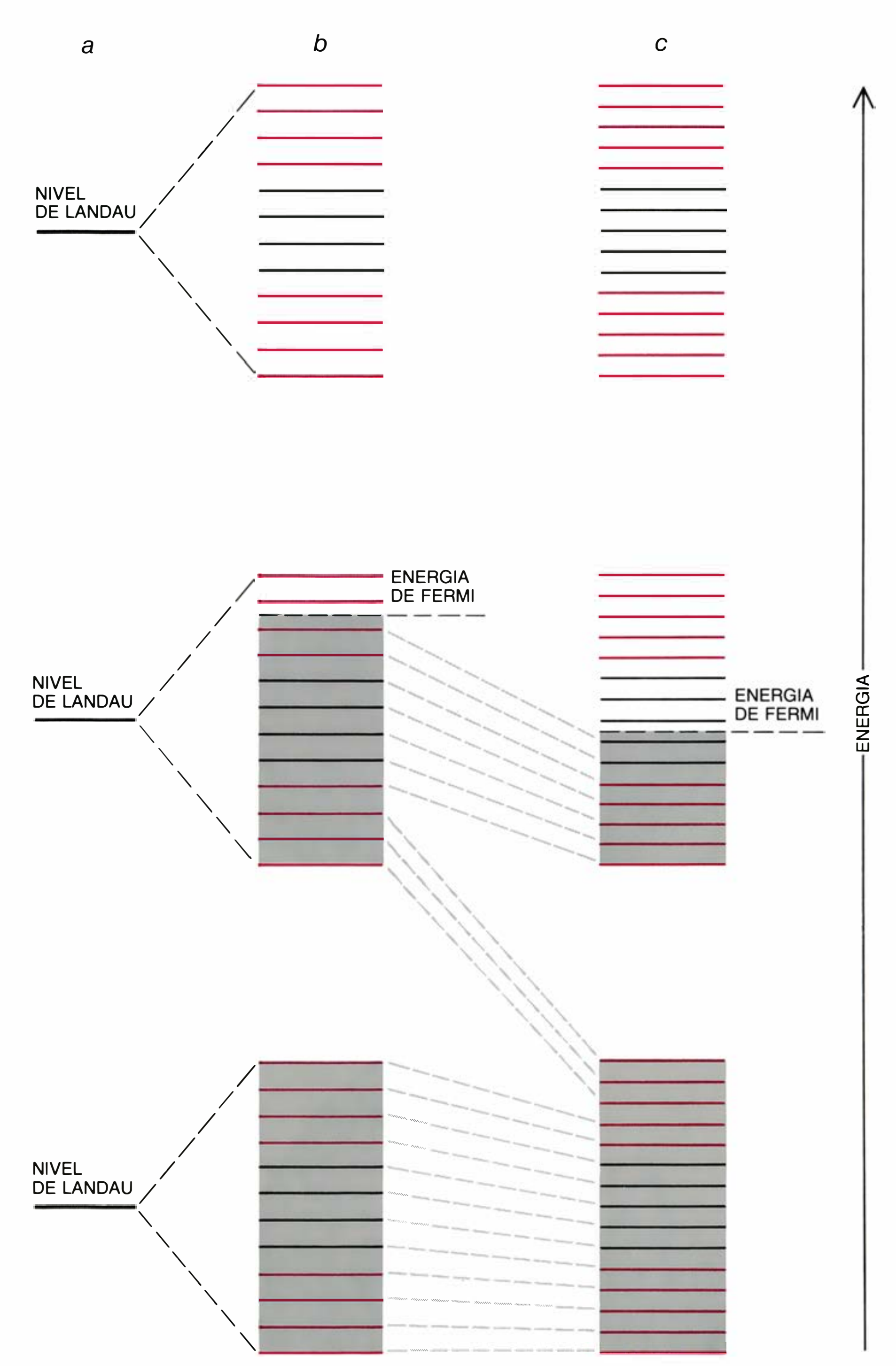
4. LAS CICLOIDES describen el movimiento de los electrones en un campo eléctrico y otro magnético, combinados. En ausencia del campo eléctrico, un electrón sometido a un campo magnético se mueve sobre una circunferencia en el plano perpendicular al campo (*a*). Cuando se aplica un campo eléctrico, el electrón dibuja una cicloide: se mueve en una circunferencia alrededor de un centro guía que procede, a su vez, sobre una línea recta. La línea que describe el centro guía es perpendicular al campo eléctrico y al campo magnético. La velocidad con que viaja el centro guía resulta ser inversamente proporcional a la intensidad del campo magnético y directamente proporcional a la intensidad del campo eléctrico; cuando el campo eléctrico se hace más intenso o el campo magnético se torna más débil (*b-d*), el movimiento de los electrones deviene entonces más lineal y el centro guía se desplaza con mayor "velocidad de deriva" (*flechas de color*).

órbitas sin que se superpongan. Resulta que la constante de proporcionalidad entre el número de estados independientes por unidad de área y la intensidad del campo magnético depende exclusivamente de la constante de Planck y de la carga del electrón.

Hasta aquí hemos dejado de lado un aspecto del sistema experimental real que reviste enorme interés en lo que concierne a todas las propiedades relacionadas con la conducción eléctrica: la influencia de las impurezas o defectos del cristal semiconductor que sostiene la delgada lámina de electrones de conducción. Las impurezas desempeñan un papel particular en la resistencia eléctrica de metales y semiconductores. Gran parte de la disipación de energía que caracteriza a una resistencia se produce cuando los electrones se desvían al colisionar con átomos de impurezas o defectos existentes en el retículo cristalino. Paradójicamente, la presencia de impurezas es lo que lleva a la desaparición de la resistencia eléctrica, así como a la aparición de mesetas ("plateaus") en la resistencia de Hall, lo que origina el efecto Hall cuántico.

En presencia de impurezas, los numerosos estados cuánticos independientes que constituyen cierto nivel de Landau no tienen ya exactamente la misma energía. En una explicación semiclásica, podríamos decir que en algunos de los estados cuánticos es ligeramente más probable encontrar el electrón cerca de un átomo de impureza, que presenta, por ejemplo, un exceso de carga positiva. (Este sería el caso si el átomo de impureza correspondiese a un elemento de los que tienden a ceder un electrón en cuanto se les encierra en el retículo de un semiconductor.) Tales estados cuánticos serían ligeramente más estables que los otros del mismo nivel de Landau y tendrían una energía ligeramente inferior. El nivel de energía único que constituye un nivel de Landau en un cristal puro está, por consiguiente, disgregado en presencia de impurezas, formando una banda compuesta por muchos niveles energéticos distintos.

Los estados cuánticos que forman parte de cada banda energética pueden dividirse en tres clases generales. Los de la parte inferior de cada banda, es decir, los estados de energía más baja, se instalan siempre en una pequeña región de la muestra. Un electrón que se encuentre en uno de estos estados nunca abandonará tal región. Estos



5. NIVELES DE ENERGIA disponibles para un electrón que se halla confinado en una lámina bidimensional inmersa en un campo magnético perpendicular. Dichos niveles determinan las propiedades eléctricas a gran escala de la lámina. En una lámina situada en un cristal ideal (*a*), un electrón sólo puede existir en estados que tienen ciertos niveles energéticos, llamados niveles de Landau. En cristales reales, que siempre contienen impurezas, cada nivel de Landau está escindido en una banda de niveles de energía (*b*); los estados del mismo nivel de Landau, que habían tenido exactamente la misma energía, poseen ahora energías ligeramente distintas porque, en algunos estados, el electrón se halla más cerca de un átomo de impureza. Los niveles de energía son de dos tipos: los estados localizados (*en color*), en los que el electrón está atrapado en las proximidades de un átomo de impureza o defecto, y los estados extendidos, en los que el electrón puede deambular por una área grande de cristal. Los estados localizados pueden presentar energías altas o bajas, dependiendo de la carga de la impureza que los produce. Unicamente los electrones que se hallan en estados extendidos pueden transportar la corriente, de modo que las propiedades eléctricas del cristal están determinadas por los electrones que se encuentran en los estados extendidos. En virtud del principio de exclusión de Pauli, cada estado puede contener, a lo sumo, un electrón. A muy bajas temperaturas, todo estado por debajo de cierta energía, la energía de Fermi, se halla ocupado (*región sombreada*); todo estado por encima del mismo, está vacío. Cuando se incrementa la intensidad del campo magnético, hay más estados en cada banda de Landau y los electrones caen (*líneas a trazos*) a estados de energía más baja (*c*). Por consiguiente, la energía de Fermi desciende a una posición más baja en la banda de Landau o a otra inferior.

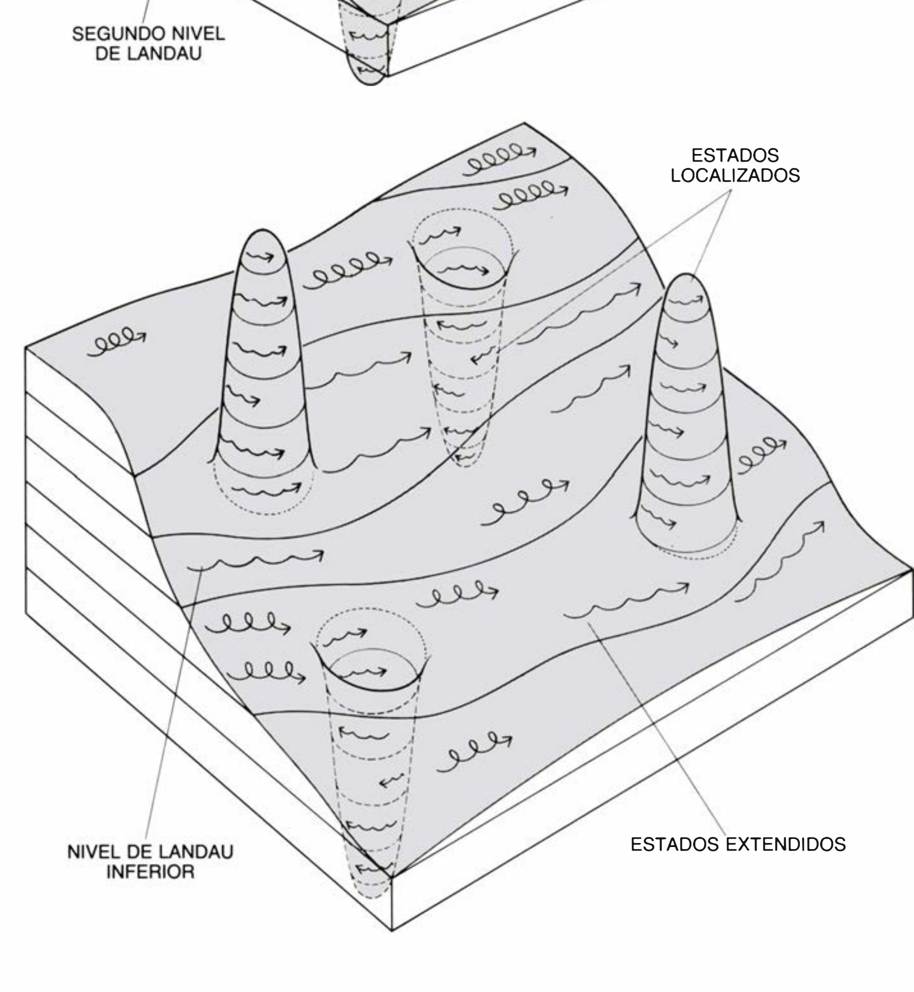
estados localizados de baja energía aparecen en "valles" de la energía potencial eléctrica, por ejemplo, en las regiones que circundan a los átomos de impureza que tienen un exceso de carga positiva. Análogamente, en la parte superior a la banda, hay estados localizados de alta energía. Se alojan cerca de los "picos" del potencial eléctrico, tal vez en las regiones que rodean a los átomos de impureza que han adquirido electrones y que tienen, por tanto, un exceso de carga negativa.

En la región central de cada banda de energía se hallan los llamados estados extendidos, cada uno de los cuales abarca una gran región del espacio. La distinción entre estados localizados y estados extendidos no es menor: los electrones que se hallan en estados localizados no pueden alejarse mucho de la posición que ocupan en la muestra ni pueden transportar la corriente [*véase la figura 6*].

Si ahora tenemos en cuenta un aspecto más del sistema experimental real, las temperaturas extremadamente bajas a las que se realizan los experimentos, será posible dar ya cierta explicación del efecto Hall cuántico entero.

A temperaturas muy bajas, los electrones de la muestra adquieren la configuración que ofrezca la energía más baja posible. De acuerdo con el principio de exclusión de Pauli, uno de los fundamentos de la mecánica cuántica, dos electrones no pueden ocupar jamás el mismo estado mecánico-cuántico. Cuando el sistema está en el nivel energético más bajo posible, todo estado cuántico disponible que se halle por debajo de cierto nivel de energía contiene exactamente un electrón, y todo estado que se halle por encima de dicho nivel no contiene ningún electrón. La energía del nivel energético ocupado más alto recibe el nombre de energía de Fermi.

Cuando hay una diferencia de potencial entre las dos caras de la muestra, no es posible definir ya una única energía de Fermi para todo el plano de electrones de conducción; la energía de Fermi varía ligeramente de un punto a otro. En cada región del espacio, los estados con energías más bajas que el nivel de Fermi local están ocupados, mientras que los estados que se hallan por encima del mismo están vacíos. El potencial medido por un voltímetro conectado a dos puntos que se hallen en caras opuestas de la muestra establece la diferencia entre las energías de Fermi locales correspondientes a los dos puntos: la diferencia entre las ener-



6. REPRESENTACION GEOMETRICA de la energía de un conjunto de electrones confinado en una lámina bidimensional sometida a un campo magnético intenso. Ilustra la diferencia entre los estados localizados y los estados extendidos y explica el efecto Hall cuántico. El eje vertical representa la energía de los electrones en dos bandas de Landau, limitadas a una pequeña sección de la muestra. Los picos y los valles en el potencial eléctrico se deben a impurezas aleatoriamente distribuidas en la muestra. Los electrones describen trayectorias cicloidales paralelas a las líneas de contorno (líneas de potencial eléctrico constante); así pues, los contornos cerrados alrededor de picos y valles representan estados localizados en los que los electrones orbitan en una pequeña región del espacio, mientras que los contornos abiertos representan estados extendidos que pueden ir de un extremo de la muestra al otro. A bajas temperaturas, los estados electrónicos cuya energía está por debajó del nivel de Fermi local (*plano de color*)

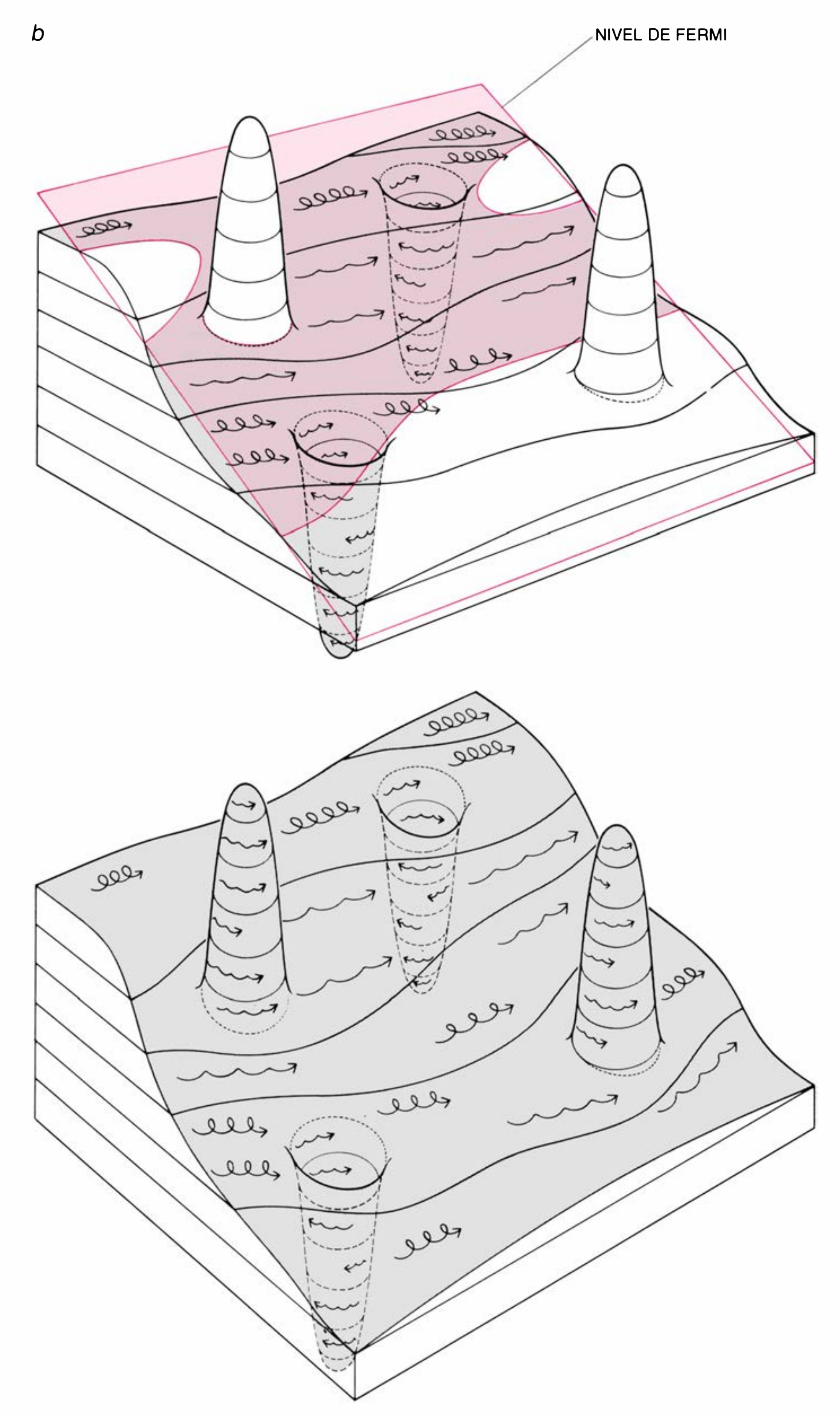
gías de Fermi locales representa la diferencia entre las energías de los electrones de máxima energía en cada punto.

Supongamos que una corriente fluye a lo largo de una muestra situada en un campo magnético perpendicular y que el nivel de Fermi del conjunto de electrones se halla en la sub-banda de estados localizados cercana al borde superior de cierta banda de Landau. En este caso, todos los estados extendidos, así como los estados localizados de baja energía de la banda de Landau, estarán ocupados y también lo estarán algunos de los estados localizados de alta energía. Supongamos, ahora, que la intensidad del campo magnético se incremente gradualmente y que, al mismo tiempo, la corriente se ajuste, sin solución de continuidad, de manera tal que la tensión de Hall entre los dos bordes permanezca constante.

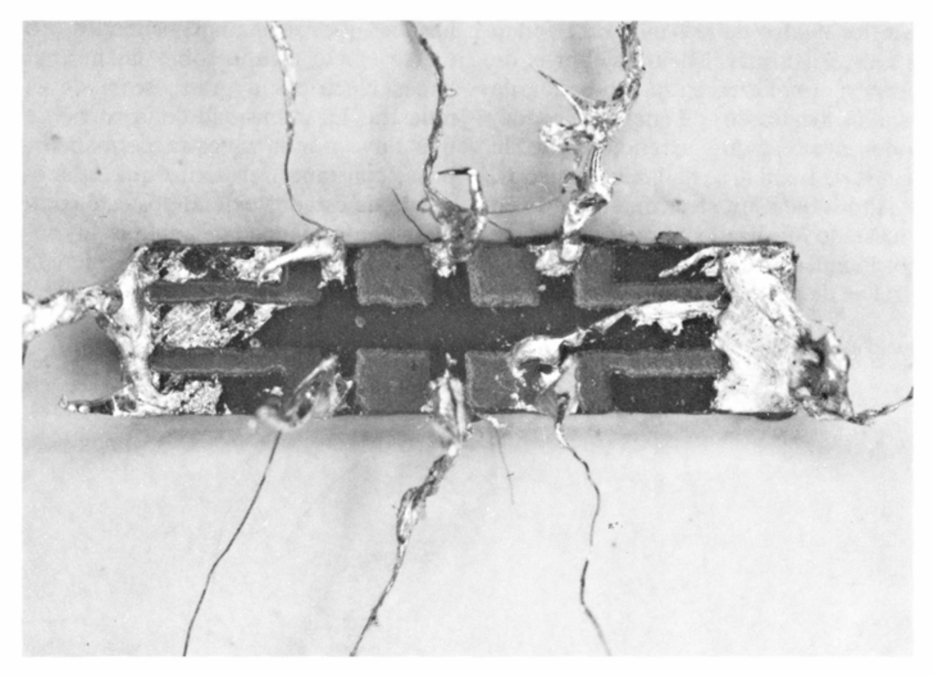
Dado que el número de estados cuánticos independientes por unidad de área es directamente proporcional al campo magnético aplicado, cuando el campo magnético aumenta el número de estados cuánticos independientes en cada nivel de Landau crecerá proporcionalmente: en toda región del espacio interior a la muestra aparecen disponibles estados cuánticos adicionales que tienen aproximadamente la misma energía que los estados vecinos.

Muchos de los nuevos estados disponibles estarán por debajo del nivel de Fermi local y, por tanto, los electrones de los estados ocupados de energía elevada bajarán a llenarlos; estos electrones provendrán, en general, de los estados localizados de alta energía cercanos al nivel de Fermi. Al quedar vacíos tales estados, el nivel de Fermi –la energía del estado ocupado más elevado– desciende a una posición inferior dentro de la banda de Landau [*véase la figura 5*]. Mientras el nivel de Fermi permanezca en la sub-banda de estados localizados de energía elevada, todos los estados extendidos de la banda de Landau se hallarán ocupados.

Ahora bien, un electrón atrapado en un estado localizado no puede moverse por la muestra; por consiguiente, la variación de la fracción de estados localizados que se hallan ocupados no ejerce efecto alguno sobre las propiedades eléctricas a gran escala de la muestra. La intensidad de la corriente que fluye por la muestra permanece, pues, constante, en tanto que la sub-banda de estados extendidos esté completamente ocupada: aunque el aumento del campo magnético decelere la velocidad de avance de todos los elec-



están llenos (*región sombreada*), mientras que los demás están vacíos. Las superficies de energía de los electrones y el nivel de Fermi local están desviados de la horizontal en razón de la diferencia de potencial (el potencial de Hall) que existe entre las dos caras de la muestra. Cuando el nivel de Fermi se halla en una región de estados localizados (*a*), se produce el efecto Hall cuántico: la resistencia de Hall (es decir, el potencial de Hall dividido por el flujo de corriente) no se ve afectada por pequeños cambios en parámetros tales como la intensidad del campo magnético o el número de electrones que hay en la capa. Es decir, un pequeño cambio en el nivel de Fermi no modificaría la resistencia de Hall porque no afectaría a los electrones de los estados extendidos, que son los que transportan la corriente. Cuando el nivel de Fermi está en una región de estados extendidos (*b*), un pequeño cambio en su altura afecta sensiblemente el flujo de corriente al vaciar o llenar algún estado extendido, y no se produce el efecto Hall.



7. DISPOSITIVO EXPERIMENTAL (una heterounión), diseñado para medir el efecto Hall cuántico. Presenta características irregulares y una geometría imprecisa; sin embargo, produce resultados que son precisos hasta aproximadamente una parte en seis millones (la misma precisión con la que se conocen las constantes fundamentales del electromagnetismo y de la mecánica cuántica). El aspecto más sorprendente del efecto Hall cuántico es que, dentro de límites muy amplios, la finísima precisión de los resultados experimentales no depende de la forma, de la pureza ni del procesado de la muestra. (Foto de H. L. Störmer.)

trones portadores de corriente, este efecto queda exactamente cancelado por el incremento del número de electrones disponibles para transportar corriente, debido a la creación de nuevos estados extendidos.

Como el potencial de Hall se está manteniendo constante, el hecho de que la corriente no cambie al variar el campo magnético implica que la resistencia de Hall (eso es, el potencial de Hall dividido por la intensidad de corriente que fluye) permanece también constante. Puede, asimismo, realizarse un experimento equivalente en el que la corriente se mantiene constante externamente mientras se mide el potencial de Hall; en este experimento, el potencial de Hall y, por tanto, la resistencia de Hall, permanecen constantes cuando el nivel de Fermi se halla en la sub-banda de estados localizados. Cuando el nivel de Fermi se encuentra en la sub-banda de estados localizados, la resistencia de Hall permanece invariable al modificarse el campo magnético. Esta es la meseta de la resistencia de Hall que caracteriza el efecto Hall cuántico.

Al ir aumentando la intensidad del campo magnético, llegará un momento en que se agotarán las existencias de electrones en los estados localizados de energía elevada, por lo que el nivel de Fermi caerá a la sub-banda de estados extendidos. Mientras el nivel de Fermi desciende por la sub-banda de estados extendidos, algunos de ellos van quedando vacíos. Como la sub-banda portadora de corriente está, entonces, sólo parcialmente ocupada, la intensidad de la corriente que fluye decrece y, por tanto, aumenta la resistencia de Hall. La resistencia de Hall seguirá creciendo si lo hace el campo magnético, en tanto el nivel de Fermi permanezca en la sub-banda de estados extendidos.

Si se sigue aumentando la intensidad del campo magnético, llegará un momento en que los estados extendidos de la banda de Landau quedarán todos vacíos, con lo que el nivel de Fermi penetrará, de nuevo, en una sub-banda de estados localizados: los estados localizados de baja energía situados en el fondo de la banda de Landau. Si existe, al menos, una banda de Landau llena por debajo del nivel de Fermi, los estados extendidos de la banda estarán en condiciones de transportar la corriente y se observará, una vez más, el efecto Hall cuántico. Sin embargo, como los estados extendidos de una banda de Landau se han vaciado por completo, el número de sub-bandas de estados ocupados ha quedado reducido en una unidad; por consiguiente, la corriente eléctrica queda reducida en la proporción del número de sub-bandas de estados extendidos ocupadas y la resistencia de Hall es correspondientemente mayor de lo que era en la meseta precedente.

En este modelo resulta fácil comprender por qué la razón entre las resistencias de Hall de dos mesetas cualesquiera debe ser igual a un cociente de enteros. El motivo reside en que, para un potencial dado cualquiera, la corriente es directamente proporcional al número de sub-bandas de estados extendidos ocupadas; y en cada meseta hay, lleno, un número entero de tales sub-bandas.

Si se sigue incrementando aún más el campo magnético, el nivel de Fermi descenderá, atravesando la región de estados localizados del fondo de una banda de Landau, hasta los estados localizados de energía elevada de la parte alta de la siguiente banda de Landau. La resistencia de Hall permanecerá constante en el valor de la nueva meseta hasta que el nivel de Fermi alcance la región central de estados extendidos de esta nueva banda de Landau.

La segunda característica sorprendente del efecto Hall cuántico es que la corriente fluye a través de la muestra sin ninguna resistencia, es decir, sin disipar energía. Esta propiedad puede explicarse mediante el principio de conservación de la energía.

Para disipar energía, un electrón debe efectuar una transición de un estado energético a otro de energía más baja, disipando el exceso de energía en forma de energía de vibración, o calor, dentro del retículo cristalino. Dos electrones no pueden ocupar el mismo estado mecánico-cuántico; por tanto, el nuevo estado en que caiga el electrón será uno que estuviera ya vacío. Si el nivel de Fermi está en la sub-banda de energías correspondiente a los estados extendidos, un electrón puede encontrarse en un estado extendido que está por debajo de la energía de Fermi local en una región del espacio, pero que alcanza otra región en la que su energía está por encima de la energía de Fermi local. En esta última región caben estados con energías intermedias entre la energía del estado del electrón y la energía de Fermi local. El electrón podría caer a uno de esos estados.

En particular, puede ocurrir que un electrón entre en la muestra por una región donde la energía de Fermi local sea alta y salga de la muestra por una región donde la energía de Fermi local sea más baja. Tal electrón podría caer a un nivel inferior mientras atraviesa la muestra, saliendo con una energía menor que la que tenía al entrar, después

de haber disipado parte de su energía en forma de calor. De este modo, la muestra ofrecería resistencia eléctrica: los electrones perderían energía al atravesarla.

Sin embargo, si el nivel de Fermi se halla entre los niveles de energía correspondientes a estados localizados, la muestra no puede ofrecer resistencia eléctrica. Los electrones de energía más elevada –aquellos cuya energía está cerca del nivel de Fermi– se alojarán todos en regiones del espacio muy pequeñas. En tales regiones, los estados desocupados tienen energías más elevadas que la energía de Fermi local. Puede haber estados vacíos de energía menor en alguna parte distante de la muestra, pero es imposible que el electrón realice una transición entre estados localizados que se hallen distantes en el espacio. Por consiguiente, los electrones no pueden caer a estados de energía más baja y, en consecuencia, no pueden disipar energía.

Esta explicación sobre la ausencia de resistencia ordinaria presupone una condición importante: que exista al menos una banda de Landau, por debajo de la energía de Fermi, que tenga los estados extendidos ocupados. Dichos estados transportan la corriente, que pasa sin disipación. Si el nivel de Fermi desciende hasta que todos los estados inferiores al mismo son localizados, no podrá fluir corriente alguna y la muestra se comportará como un aislante perfecto, no como un conductor perfecto.

Los estados localizados creados por la presencia de átomos de impurezas y defectos revisten una importancia crucial en esta explicación del efecto Hall cuántico. Los estados localizados actúan como depósitos de electrones; así, para cierto intervalo de valores de la intensidad del campo magnético, los estados extendidos de cada banda de Landau están del todo vacíos o completamente ocupados. Sin estos depósitos, la resistencia únicamente se anularía cuando hubiera exactamente el número de electrones suficiente para llenar un nivel de Landau; no se formarían, entonces, mesetas de ningún tipo en la resistencia de Hall.

Los argumentos que se han dado hasta aquí tejen una explicación plausible del efecto Hall cuántico entero. Mayor dificultad entraña justificar el efecto Hall cuántico fraccionario. El efecto entero admite una representación plástica inmediata, con el número entero de niveles de Landau ocupados. El efecto fraccionario, por contra, aparece únicamente cuando un nivel de Landau está parcialmente ocupado. Por ejemplo, una meseta donde el recíproco de la resistencia de Hall sea igual a un tercio de la carga del electrón al cuadrado dividida por la constante de Planck se observa cuando el nivel de Landau más bajo está lleno en su tercera parte, aproximadamente.

Ni siquiera un tratamiento mecánico-cuántico completo del modelo que se ha seguido hasta este momento –el de electrones independientes que se mueven bajo la influencia de campos eléctricos y magnéticos– muestra tipo alguno de estabilidad especial cuando se halla ocupada sólo una fracción de los estados disponibles. Para explicar el efecto Hall cuántico fraccionario, hemos de tener en cuenta las interacciones mutuas de los electrones; ello significa que debemos utilizar una función de onda mecánico-cuántica que dependa simultáneamente de las posiciones de todas las partículas del sistema.

El avance más importante en la comprensión del efecto Hall cuántico fraccionario se produjo en 1983, cuando Robert Laughlin, del laboratorio estadounidense Lawrence Livermore, construyó una función de onda que presentaba la estabilidad requerida cuando la fracción de estados ocupados adquiría los siguientes valores: 1/3, 1/5, 1/7, 2/3, 4/5 o 6/7. Cada una de estas fracciones es igual a la unidad dividido por un entero o impar, o igual a la unidad menos tal fracción. La importancia de este punto queda superada por el hecho de que las funciones de onda construidas para explicar el efecto Hall cuántico fraccionario difieren radicalmente de cualquier otra función de onda encontrada antes en física.

El hecho de que el denominador sea impar en cada fracción estable obedece, en parte, a una condición mecánico-cuántica que, relacionada con el principio de exclusión de Pauli, concierne, empero, a una componente de la función de onda conocida con el nombre de fase; ésta carece de análogo clásico. Tras el trabajo original de Laughlin, otros físicos teóricos, entre ellos el propio Laughlin, Duncan Haldane, que a la sazón estaba en la Universidad del Sur de California, y el propio autor, descubrieron funciones de onda que eran estables para otras fracciones con denominadores impares, tales como 2/5, 2/7 y 3/7.

Una predicción teórica importante, consecuencia del análisis de Laughlin, es la siguiente: si se añade un electrón extra a un sistema cuyo nivel de Landau está ya lleno en su tercera parte, la carga extra debería aparecer en tres lugares separados dentro de la muestra. En cada una de esas posiciones debería aparecer, exactamente, un tercio de la carga del electrón. Cargas fraccionarias similares deberían aparecer si se añaden o si se quitan electrones en cualquiera de las mesetas ("plateaus") del efecto Hall cuántico fraccionario. Las cargas fraccionarias reciben el nombre de cuasipartículas; se supone que se comportan de manera muy semejante a las partículas cargadas.

En el efecto Hall cuántico fraccionario, son los estados de las cuasipartículas de carga fraccionaria, y no los de los electrones, los que están localizados en las cimas y valles del potencial que existe en la región circundante a los átomos de impurezas o defectos. Los estados localizados actúan, entonces, como depósitos que mantienen la resistencia de Hall constante para un intervalo finito de valores de la intensidad del campo magnético.

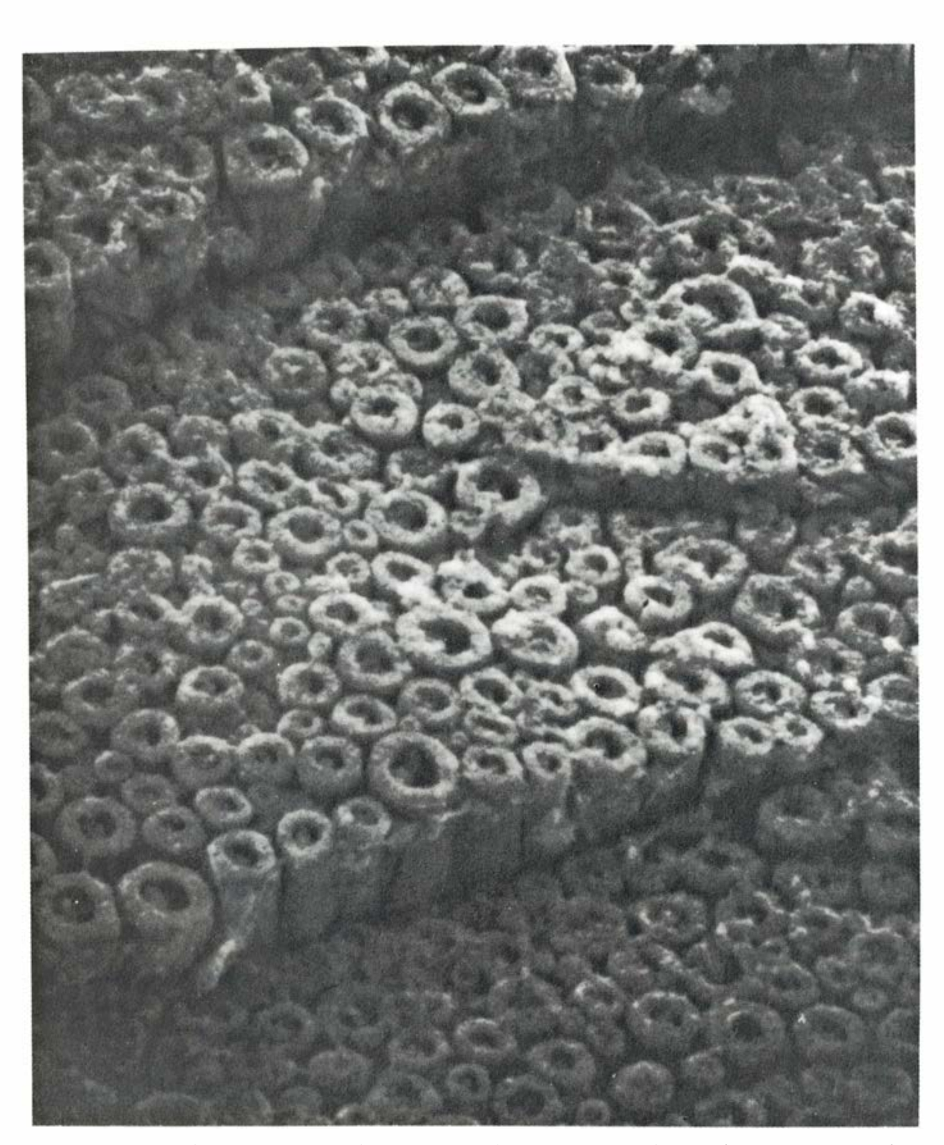
Parece inverosímil que el efecto Hall cuántico pueda emplearse para construir un aparato que revolucione la economía. La exigencia de campos magnéticos intensos y temperaturas bajísimas añade unos gastos que excluyen cualquier aplicación práctica. Sin embargo, los investigadores harán sin duda uso de las mesetas de Hall en el laboratorio como patrones precisos de resistencia elétrica; además, el efecto Hall cuántico se incorporará probablemente a experimentos destinados a establecer valores más precisos para las constantes fundamentales de la mecánica cuántica, la electricidad y el magnetismo. Cuando se los combine con otras mediciones, estos experimentos pueden proporcionar tests más rigurosos de la teoría fundamental de las interacciones electromagnéticas, completando los tests de esta teoría que se están llevando a cabo con aceleradores de altas energías.

La verdadera importancia del efecto Hall cuántico no radica, sin embargo, en ninguna de estas aplicaciones, sino en la perspectiva, completamente nueva, que los físicos han adquirido acerca de las peculiares propiedades de los sistemas de electrones sometidos a campos magnéticos intensos y acerca de las regularidades encubiertas derivadas de las leyes matemáticas de la mecánica cuántica. No menos importante: este efecto nos recuerda que la naturaleza puede encerrar otros sorprendentes estados de la materia de los que no tenemos ni sospecha.

Ciencia y sociedad

Reunión de 175.000 neuritas

En nuestro organismo, los nervios constituyen un sistema comparable a la red telefónica. Son sus hilos las neuritas, las prolongaciones de las neuronas, miles y miles de filamentos de tamaño microscópico. Agrupadas en estructuras que semejan cables (los nervios) conectan el cuerpo entero con el conmutador central: la médula espinal y el cerebro. Cada neurita, con cubierta aislante, está encerrada en un delgado tubo. Este, de origen membranoso (la lámina basal), mide 70 nanómetros de espesor y se extiende sin solución de un extremo a otro del nervio. Nunca se le dio mayor importancia a esas membranas, que no son celulares, si bien constituyen el microesqueleto del nervio y cumplen una importante función en la regeneración del nervio lesionado. Cuando se produce algún daño, y en la medida en que sigan intactas las células de las que parten, las fibras nerviosas pueden regenerarse a sí mismas. Así lo demostró magistralmente Santiago Ramón y Cajal, quien describió el mecanismo de esa reparación: se desprende la porción de las neuritas situadas bajo la zona afectada, al tiempo que comienzan a crecer y alargarse los segmentos que siguen conectados a los centros. Los renuevos invaden rápidamente la parte dañada del nervio en un irresistible movimiento dirigido a restablecer el contacto con la periferia. De no lograrse una conexión correcta, el órgano o miembro afectado quedará inmóvil o insensible, tan inútil como un teléfono desconectado.



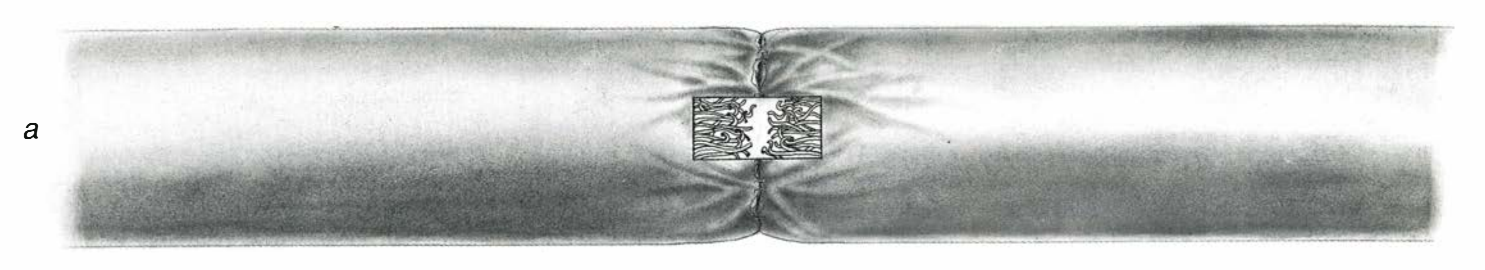
1. Fotografía de una sección de un nervio humano tomada en el microscopio electrónico de barrido. Se ven unas 250 neuritas, envuelta cada una en su vaina. La técnica de fijación empleada ha inmovilizado las neuritas blandas sin afectar su orden normal. La imagen pone de relieve la abundancia y diversidad de elementos que han de ponerse en conexión para restablecer el funcionamiento del nervio.

Las lesiones nerviosas se agrupan en dos categorías. En unos casos sólo afectan al centro, frágil, de las fibras; así sucede cuando se aplasta el nervio sometido a una presión local excesiva. No se rompe la continuidad de la capa envolvente: observado desde el exterior, el nervio parece normal. En el interior permanecen intactos los tubos de lámina basal, de sorprendente resistencia si atendemos a su escaso grosor; servirán de guía que ayude a orientarse a las neuritas en su regeneración. Las fibras se deslizan por esos tubos de membrana basal cual si éstos fueran túneles, cruzan la porción dañada del nervio y se dirigen sin error posible hacia su destino periférico correcto. Por lo común, esa guía permite una excelente restitución del nervio y la recuperación satisfactoria sin necesidad de operar.

Las consecuencias del segundo tipo de lesión, la sección completa del nervio, son mucho más graves. En la zona dañada se destruyen todas las sendas y se pierde el efecto de guía. Los renuevos, brotados en la rama superior, crecerán erráticamente por la región del corte. Alcanzada la rama inferior del nervio, penetran en las estructuras ilesas y siguen a ciegas y sin detenerse la primera guía que encuentren. El punto de sutura constituye, por tanto, una zona de importancia decisiva. El menor error de reparación puede enviar una neurita motora por la senda de otra sensorial, o una neurita sensorial por el recorrido de otra motora. Esas fibras quedan definitivamente inutilizadas. Cabe también que una neurita motora de un músculo flexor voluntario se dirija a un músculo extensor o a uno involuntario. Resultan tan patentes las consecuencias de tal enredo que difícilmente se logra en ese caso la recuperación de la función.

De hecho, el promedio de restablecimientos funcionales de las secciones es muy bajo. Lo deseable sería que, al reparar el nervio, el cirujano dejara cada neurita frente a su senda correspondiente; desgraciadamente, las técnicas microquirúrgicas más precisas de que hoy se dispone ni siquiera anhelan ese objetivo. Un nervio típico de micromamífero como la rata suele poseer

a



b

c

d

e

f

*2. La zona de sutura de un nervio se halla sujeta a una gran tensión. Esta, irregular en la vaina, desorganiza las fibras (*a*). En los dibujos subsiguientes se ilustra el principio de de Saint-Venant, relativo a la distribución de fuerzas en un cuerpo elástico. En el caso de cilindros sometidos a tensión axial, la distribución de las tensiones depende de cómo se aplique la fuerza, pero sólo en los extremos (*b*); en las secciones transversales alejadas de los extremos, la tensión es casi uniforme (*c*). Si se ejerciera una tensión a cierta distancia del borde, no habría tensión en el extremo (*d*), sino un desplazamiento uniforme de toda la estructura (*e*). Resulta, pues, conveniente conseguir que en la intervención microquirúrgica toda la tensión mecánica se produzca lateralmente, lejos de la zona de reparación (*f*).*

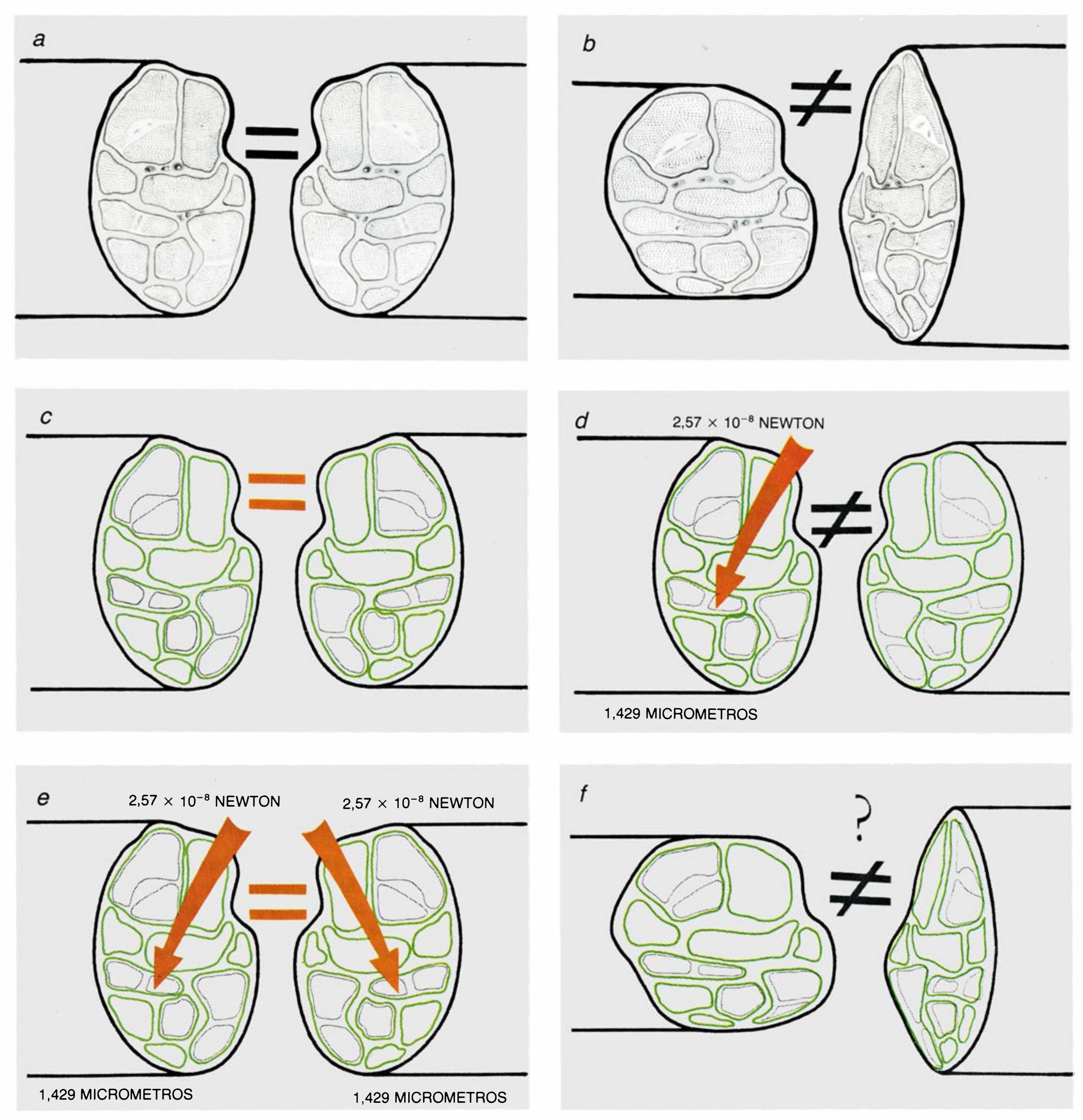
unas 3000 fibras; nuestro nervio ciático contiene hasta 175.000 neuritas. Puesto que su diámetro mide entre 0,5 y 30 micrometros, parece empresa imposible conectarlas correctamente [*véase la figura 1*]. Hoy el cirujano se limita a alinear lo mejor que puede las dos ramas del nervio. Por supuesto, las neuritas regeneradas crecen entonces en total desorden y se entrelazan de forma arbitraria; de los malos resultados que cabe esperar de ello suele culparse a la madre naturaleza. Sin embargo, existe una solución sencilla para el apareamiento preciso de un número indefinido de fibras.

Por lógica, basta con satisfacer dos condiciones: en primer lugar, que las fibras y los tubos de lámina basal se mantengan en correcta alineación longitudinal en el interior de las ramas

nerviosas; debe evitarse, por ejemplo, que formen espirales. En segundo lugar, que la alineación transversal sea perfecta: deben disponerse las neuritas frente a su senda correspondiente en la rama opuesta.

Parece extremadamente difícil mantener la alineación longitudinal. Las fibras nerviosas son tan blandas que la mínima presión modifica la dirección en que miran. En los nervios ilesos la continuidad de las vainas impide la desorganización, pero en la punta del nervio seccionado las neuritas se mueven, se enroscan y se doblan cual algas marinas. La sutura de las ramas nerviosas empeora considerablemente la situación; por muy delicados que sean los puntos, impone una tensión irregular al nervio. En los puntos mismos se genera tracción, mientras que entre ellos se crea distensión. La irregularidad de la tensión que se ejerce sobre la vaina desencadena una profunda desorganización longitudinal de las fibras en los extremos de las ramas.

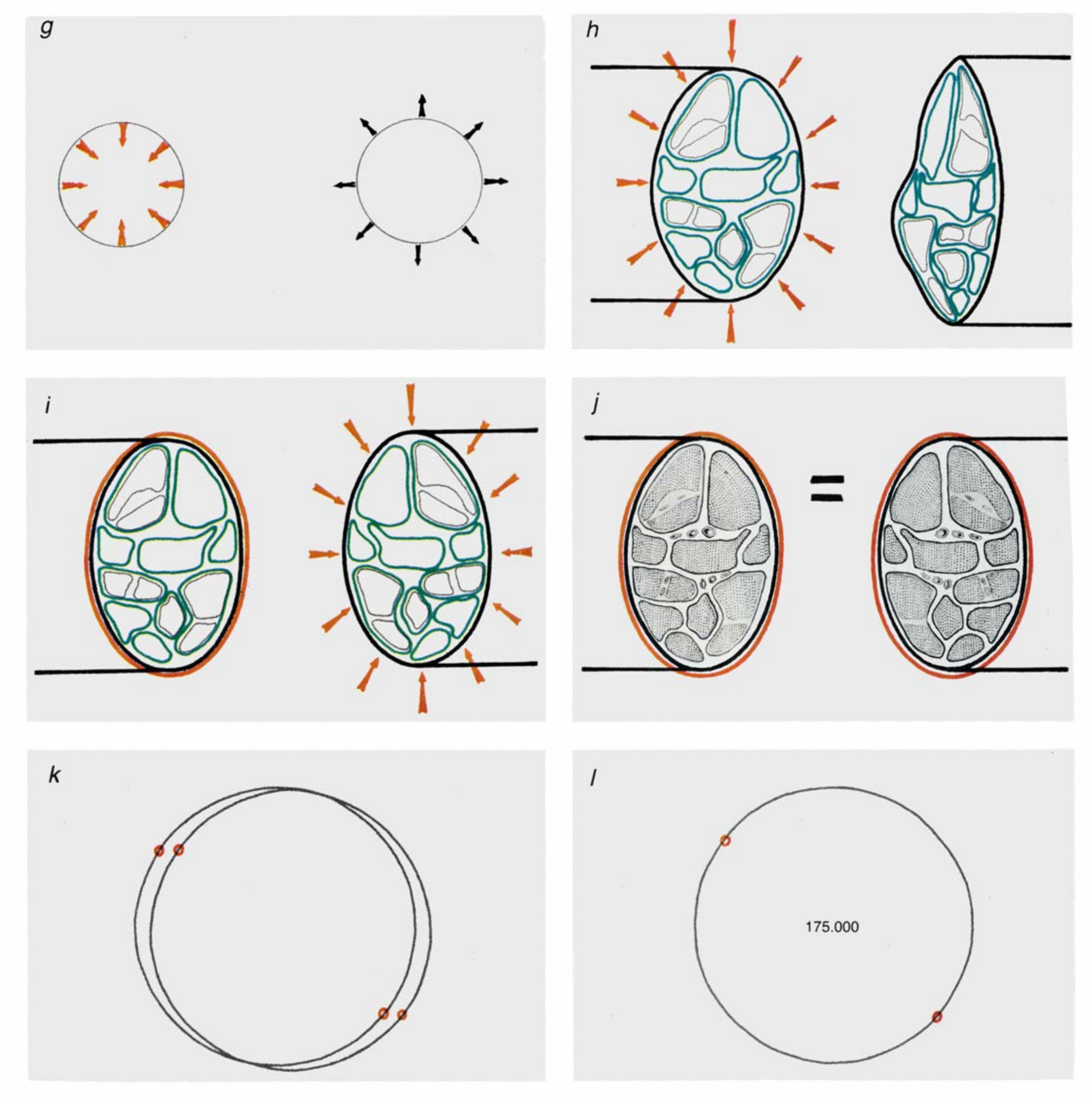
El principio de de Saint-Venant, relativo a la distribución de las fuerzas de tensión en los cuerpos elásticos, aporta una solución sencilla. Para el caso de los cilindros elásticos sometidos a tensión axial, el principio establece que la distribución de las tensiones depende de cómo se aplique la fuerza tensora, pero sólo en las proximidades de los extremos que registran la carga; en las



*3. Cuando se secciona un nervio, los extremos o cabos de las dos ramas son iguales (*a*). Ahora bien, la blandura del tejido posibilita la deformación arbitraria y la pérdida de la mutua semejanza (*b*). Ambos cabos comparten asimismo otras características físicas: resistencia mecánica y elasticidad (*c*), de suerte que si se aplica una presión a un punto de una de las ramas se produce un desplazamiento correspondiente de las estructuras (*d*). Si ejerciéramos una fuerza equivalente en un punto simétrico del otro cabo, observaríamos un desplazamiento parejo (*e*). ¿Hasta dónde llega esa simetría? Nos lo dirá la deter-*

secciones transversales alejadas de los extremos, la distribución de la tensión es casi uniforme. En la práctica, una distancia equivalente a 1,5 o 2 diámetros de la estructura basta para apreciar ese efecto, enunciado por el ingeniero civil A. de Saint-Venant en una larga "Memoria sobre la torsión de los prismas" recogida en la edición de 1856 de las Memorias de la Academia de Ciencias de París. Pese a que el principio se ha demostrado válido siempre que se ha sometido a prueba, un siglo después de su enunciado sigue sin recibir explicación completa. Pero podemos inferir un postulado: si se invirtiera la situación, esto es, si se aplicara la tracción a cierta distancia del borde, no deberían registrarse tensiones en el extremo de la rama, antes bien, la estructura entera debería desplazarse uniformemente.

El postulado sugiere un sencillo método de reunión de las ramas nerviosas. Se adhiere a la vaina externa de la rama, con la tensión apropiada y a cierta distancia del extremo, una pequeña lámina de cualquier materia tolerable por el organismo; se ejercen así lateralmente y lejos de la zona de reparación todas las tensiones mecánicas. Seguidamente se aproximan con suavidad las ramas nerviosas. El análisis histológico de ese tipo de reuniones muestra siempre una espectacular diferencia respecto de las obtenidas por



minación de la fuerza global que, aplicada en la zona externa del nervio, producirá los desplazamientos deseados (f); y lo mejor es optar por una fuerza circular centrífuga o centrípeta (g). Las deformaciones internas, aunque impredecibles, no son arbitrarias (h): las mismas que resultan en el otro cabo si procedemos igual (i). Se restablece luego la simetría perdida (j). Cuando se ajustan dos puntos de la circunferencia (k), quedan apareadas, en teoría, una a una, las 175.000 neuritas (l). En la práctica clínica bastaría con la correcta alineación de familias de neuritas.

el método de microsutura regular. Las fibras mantienen una alineación perfecta; no se tuercen ni se doblan. La hendidura de separación resulta igualmente satisfactoria: muy pequeña y uniforme.

Resuelta, por tanto, la primera parte del problema, se logra el alineamiento adecuado sin más que aplicar un segundo postulado.

Al seccionarse un nervio, los extremos de las dos ramas son, por definición, rigurosamente idénticos: uno imagen especular perfecta del otro [*véase la figura 3*]. Se advierte en la superficie de corte una delicada red de divisiones. En el momento justo de la sección, las dos ramas presentan una estructura exactamente simétrica; sin embargo, por ser blando el tejido nervioso, los extremos se deforman estocásticamente y suele perderse esa semejanza.

De hecho, la perfecta concordancia inicial no sólo se refiere a la morfología del nervio, sino también a las características físicas y, por tanto, a la resistencia mecánica y la elasticidad de todas las porciones. Si se aplica una presión de cierta intensidad a un punto dado de una de las ramas, las estructuras sufren el correspondiente desplazamiento; de aplicarse una fuerza igual en el punto simétrico de la otra rama, se registraría en ésta un desplazamiento equivalente. Por supuesto, las neuritas siguen el desplazamiento de las estructuras envolventes.

Las neuritas carecen de libertad de movimiento en el interior del nervio; puede aventurarse, pues, una precisión extrema de la simetría de esos movimientos en ambas ramas. Las fibras se encuentran inmersas en una red tenue, blanda y ubicua de tejido conjuntivo, una matriz gelatinosa, que apenas se distingue al microscopio, ni siquiera a gran aumento, y que determina la posición exacta que ocupa cada fibra respecto de sus vecinas. No es necesario, por tanto, intervenir sobre cada una de las neuritas; basta con determinar la fuerza global que, aplicada a la parte exterior del nervio, produce el desplazamiento deseado.

La solución más práctica consiste en ejercer una fuerza circular, positiva o negativa, con lo que se nivelan las prominencias y hendiduras de la superficie de corte. Resulta del todo impredecible la deformación interna que así se genera; en todo caso, no es aleatoria. Al someter la otra rama al mismo procedimiento de circularización, sí es predecible qué deformación experimentará su contenido interno, puesto que coincidirá exactamente con la registrada en la primera rama: se ha recuperado la simetría perdida. Al ajustarse dos puntos de la circunferencia, teóricamente se aparean una a una las 175.000 neuritas. La verdad es que en la práctica clínica bastaría con la correcta alineación de familias enteras de neuritas.

En segmentos del nervio ciático puede efectuarse una demostración indirecta de la verosimilitud de esa precisión que postula la teoría. Se toman microfotografías de la superficie del corte antes y después de aplicarse la fuerza circular. La superposición de las imágenes de ambos extremos, a cien aumentos, muestra que no pueden alinearse antes de la aplicación de la fuerza. Después de ésta, las fotografías coinciden con tal precisión que parecen una sola imagen. A ese aumento no se aprecian las neuritas, pero si se las viera tampoco habría forma de distinguir unas de otras. Ello no obstante, todos los detalles de las dos imágenes (divisiones, vasos sanguíneos o tejido conjuntivo) resultan perfectamente simétricos. También deben serlo las neuritas.

Un punto débil manifiesto del método es que no siempre provocan las lesiones un corte perpendicular limpio. Se soslaya ese problema extirpando un trozo de nervio (a menos que el fragmento eliminado sea considerable, no se perderá de forma significativa la similitud). Por regla general, puede resecarse en ambas ramas una longitud equivalente al diámetro sin que, aparentemente, se modifiquen las estructuras y disposición internas. En la mayoría de los casos, basta esa distancia para obtener superficies perpendiculares limpias.

El correcto apareamiento de las neuritas no satisface, por supuesto, todos los requisitos que exige la neurocirugía periférica para lograr el éxito. Imaginemos que el electricista ha conseguido situar todos los hilos del cable exactamente frente a los extremos que les corresponden; debe aún conectarlos, y son muchos los fallos que pueden darse en esa operación. En todo caso, sí se habrá reducido significativamente el número de posibles errores. De igual manera, de nada sirve la correcta alineación de las neuritas si no se cumplen luego otras condiciones. Constituye, sin embargo, un paso indispensable hacia la solución de los problemas que plantean las lesiones del sistema nervioso. (Luis de Medinaceli.)

Ritmos internos en la emigración de las aves

Las aves migradoras poseen un reloj interno que marcha al compás de ritmos circanuales y les informa de cuándo empezar y cuándo concluir el vuelo. Es más, les ayuda incluso en la localización del punto de destino

Eberhard Gwinner

Rara vez deja de maravillarnos la contemplación de una bandada de aves que emigran, hacia el sur en otoño o con rumbo norte en primavera. Tanto si se trata de la ordenada formación en V del barnaclas del Canadá como del avance atropellado de un grupo de estorninos, la impresión que transmiten al observador es que les mueve un poderoso impulso interno. Las aves no titubean; vuelan seguras, inequívocamente orientadas hacia una meta que queda muy lejos del alcance de nuestra vista. ¿De dónde surge el impulso que las guía hacia climas más templados en invierno y las devuelve, en primavera, a sus territorios de cría septentrionales?

Caben dos respuestas. La primera es que procede de factores externos al ave misma, factores ambientales. Entre los que acuden de inmediato a la memoria se cuentan los cambios de temperatura y del número de horas de luz por día entre una estación y otra. La disminución de la temperatura o de la lontigud del día podría bastar para desencadenar las reacciones fisiológicas que ponen al pájaro en camino hacia el sur; el incremento de esos factores instaría los cambios fisiológicos inversos, devolviendo al ave hacia el norte. El segundo tipo de explicación general es que el impulso procede del animal.

En el pasado, la mayoría de las investigaciones se centraron en la influencia de factores externos. Durante los últimos 20 años, sin embargo, mis colegas y yo hemos estudiado el efecto de los ritmos internos, trabajos que nos han llevado a concluir que ciertos ritmos endógenos, de un período aproximado de un año, ejercen un fuerte influjo sobre cuándo y cómo deben las aves empezar a moverse. Esos ciclos "circanuales" parecen constituir un marco general de la regulación temporal de la migración. Además, afectan a los detalles del vuelo migratorio, ayudando quizás a las aves a navegar hacia metas específicas. El aprendizaje y los factores ambientales externos podrían modificar los ritmos circanuales. Con todo, su identificación parece habernos mostrado un componente esencial del mecanismo que controla la migración aviar.

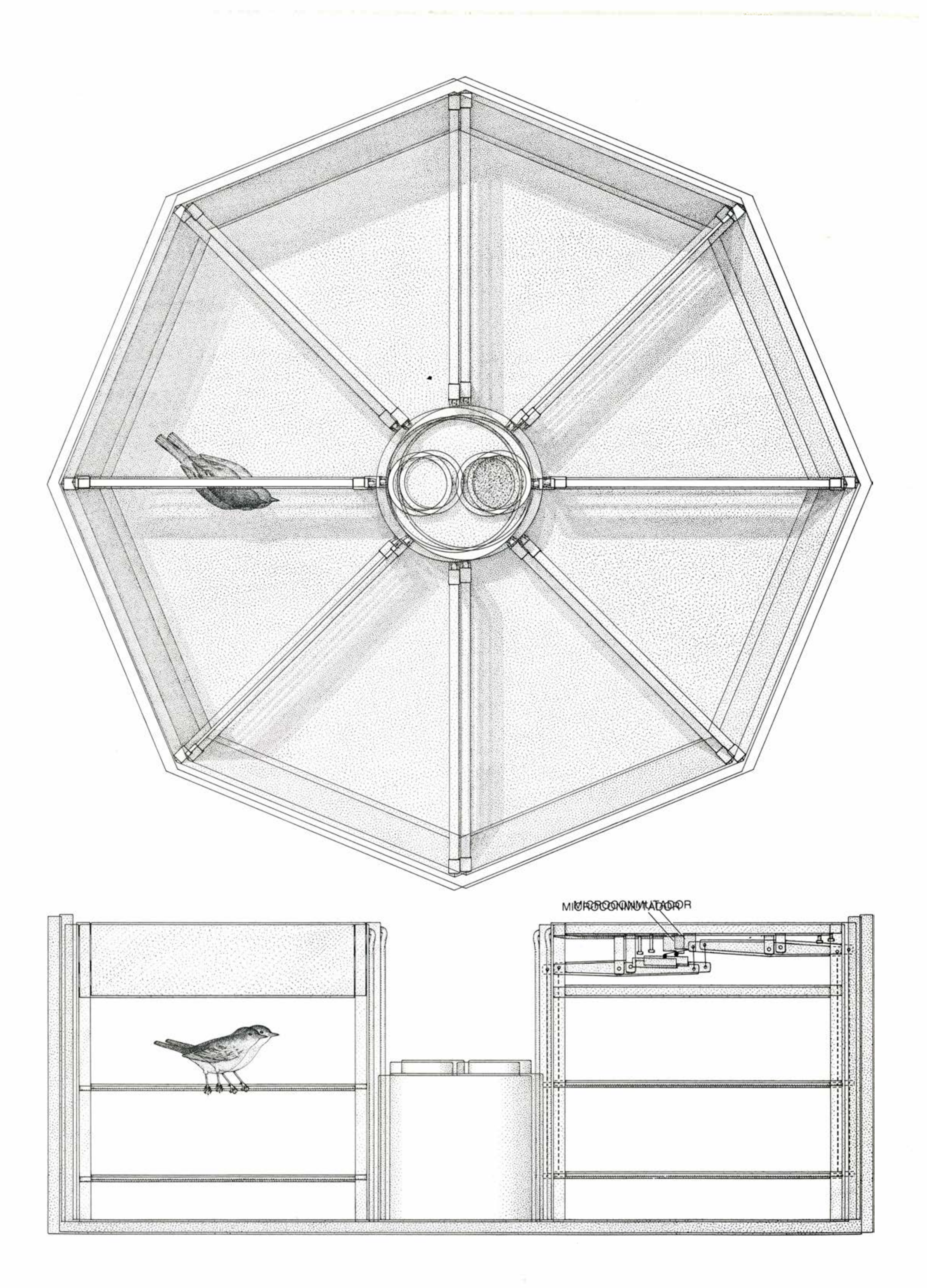
El instituto donde trabajamos, el Vogelwarte Radolfzell, se encuentra en el sur de Alemania; nuestro estudio, por tanto, se centró en especies que forman parte del sistema de emigración paleártico-africano. Constituye éste uno de los grandes sistemas de migración aviar, cada uno de los cuales comprende las trayectorias de vuelo de muchas especies. Un segundo sistema de migración une Norteamérica con América Central y del Sur; un tercero se extiende desde el noreste de Asia al sureste asiático y Australia. Dado que una gran proporción de las aves que habitan en regiones templadas emigran todos los años a tierras meridionales, esos sistemas migratorios pueden resultar muy amplios y complejos. De todos ellos, el paleártico-africano es el que más se ha estudiado.

Gran parte de lo que sabemos del sistema paleártico-africano procede de los experimentos de anillamiento que se emprendieron en Europa hacia comienzos de siglo. Desde entonces, se han anillado en Europa más de 60 millones de aves; la recuperación de más de un millón de ellas nos ha informado ampliamente sobre sus rutas y territorios de invernada. Basándose parcialmente en esos conocimientos, R. E. Moreau, del Instituto inglés Edward Grey, ha calculado que cada año invaden Africa más de cinco mil millones de aves procedentes de Europa. Los territorios de invernada de los invasores se extienden desde el norte africano hasta el cabo de Buena Esperanza, siendo muy diversas las distancias que cubren. Mientras que una chocha perdiz quizá sólo recorra unos centenares de kilómetros a través del Mediterráneo, un combatiente de Siberia tendrá que volar más de 12.000 kilómetros por Asia y Europa oriental hasta alcanzar sus cuarteles de invierno en el Africa Central.

En su mayor parte, nuestros experimentos se han practicado con currucas (familia Sylviidae) y papamoscas (familia Muscicapidae). El estudio del mecanismo de control de la emigración en aves de esas especies planteaba tres interrogantes fundamentales. El primero era cómo se controla el compás de la emigración. El segundo, cómo consigue el ave navegar hacia una meta concreta en cada etapa de la emigración y, el tercero, cómo soporta el organismo la fuerte demanda energética que impone tan prolongado vuelo, en particular cuando salva océanos o desiertos, donde el ave no puede reponer con facilidad sus agotadas reservas energéticas.

Las tres preguntas tienen interés,

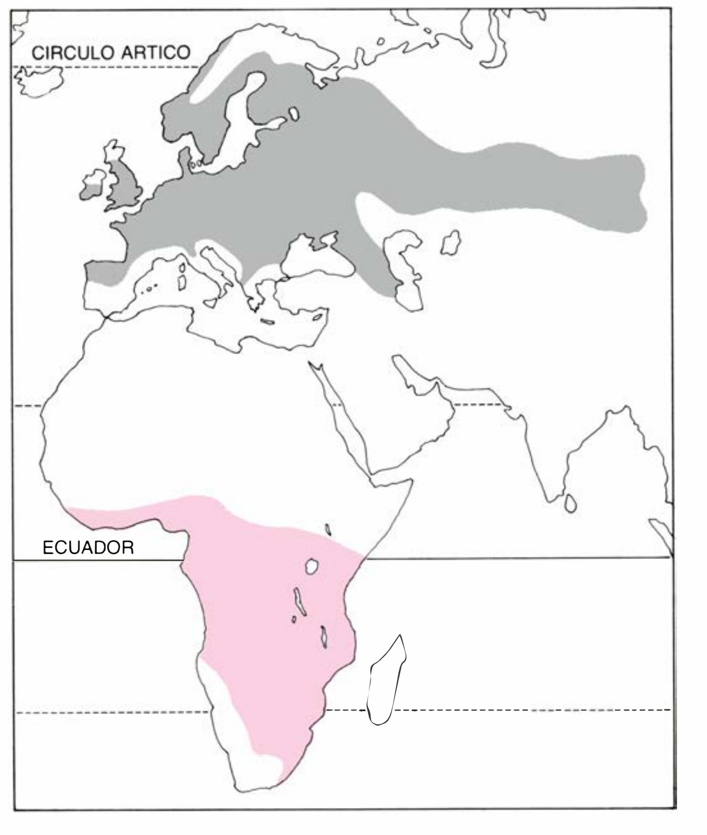
1. JAULA DE ORIENTACION empleada por el autor para comprobar las preferencias de las aves migradoras en diversas épocas del año. La jaula tiene ocho pares de perchas (*arriba*), cada una de ellas unida por un hilo a un microconmutador (*abajo*). Cuando el ave salta a una de las perchas, el microconmutador cierra un circuito y un registrador acusa los impulsos eléctricos. El aparato permite averiguar qué parte de la jaula (y por tanto, qué dirección de la brújula) prefiere el ave. El autor pudo así demostrar que las preferencias de dirección de las aves migradoras experimentan cambios instados por ritmos endógenos. El ave estudiada en los experimentos del autor era la curruca mosquitera (*Sylvia borin*), y tal es la especie mostrada.

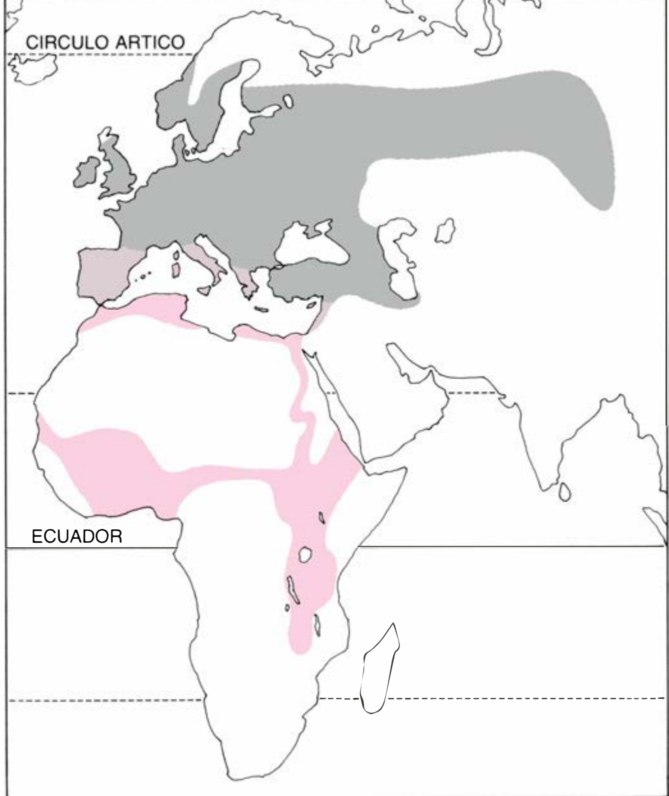

MICROCONMUTADOR

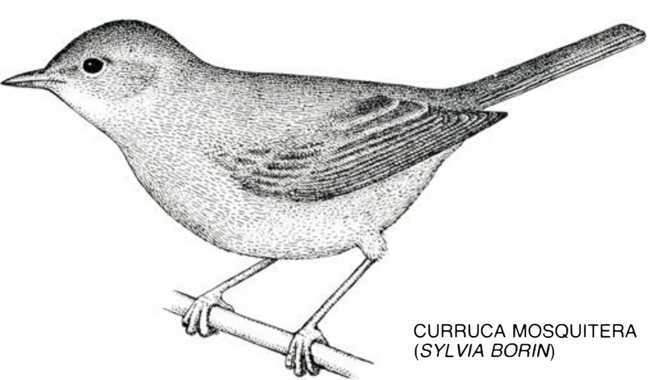
pero abordarlas a un tiempo hubiera constituido un proyecto hasta cierto punto irrealizable. Así pues, nuestra estrategia de investigación no prestó igual atención a todas ellas. Antes bien, desde el comienzo mismo de nuestro trabajo, a mediados de la década de 1960, nos concentramos en el compás de la emigración, en su control temporal. Concretamente, pretendíamos identificar el origen de la información que impulsa al ave a prepararse para emigrar, a empezar el viaje y, por último, a concluir el vuelo migratorio. Los resultados obtenidos en ese empeño por esclarecer el problema del compás quizás ayuden a responder las otras dos preguntas.

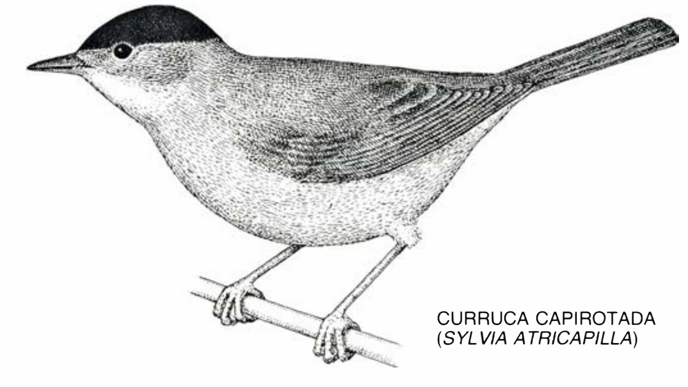
Acometimos, pues, una larga serie de experimentos encaminados a descubrir el control del compás migratorio. En muchos de ellos sometíamos aves enjauladas a una cuidadosa observación para detectar comportamientos asociados con la migración. Tales comportamientos resultaban de la mayor importancia para nuestro trabajo, porque servían de indicadores de las sucesivas etapas del ciclo migratorio anual. Ejemplo de ello es la "desazón migratoria". Las currucas y los papamoscas de nuestro estudio son migradores nocturnos: vuelan de noche y reposan de día. Si se enjaulan las aves en primavera u otoño, cuando en circunstancias normales estarían emigrando, muestran una intensa actividad nocturna. Tal desazón, que puede adoptar diversas formas, constituye un síntoma de que han comenzado los acontecimientos fisiológicos subyacentes a la migración.

La desazón migratoria desempeñó un papel fundamental en el primer experimento de la serie, realizado en 1966. Nos proponíamos averiguar cómo se inicia la migración en el mosquitero musical (*Phylloscopus trochilus*), un ave pequeña, que pasa inadvertida. Su área de reproducción se extiende por buena parte del centro y norte de Eurasia, y su territorio de invernada lo constituyen las regiones tro-







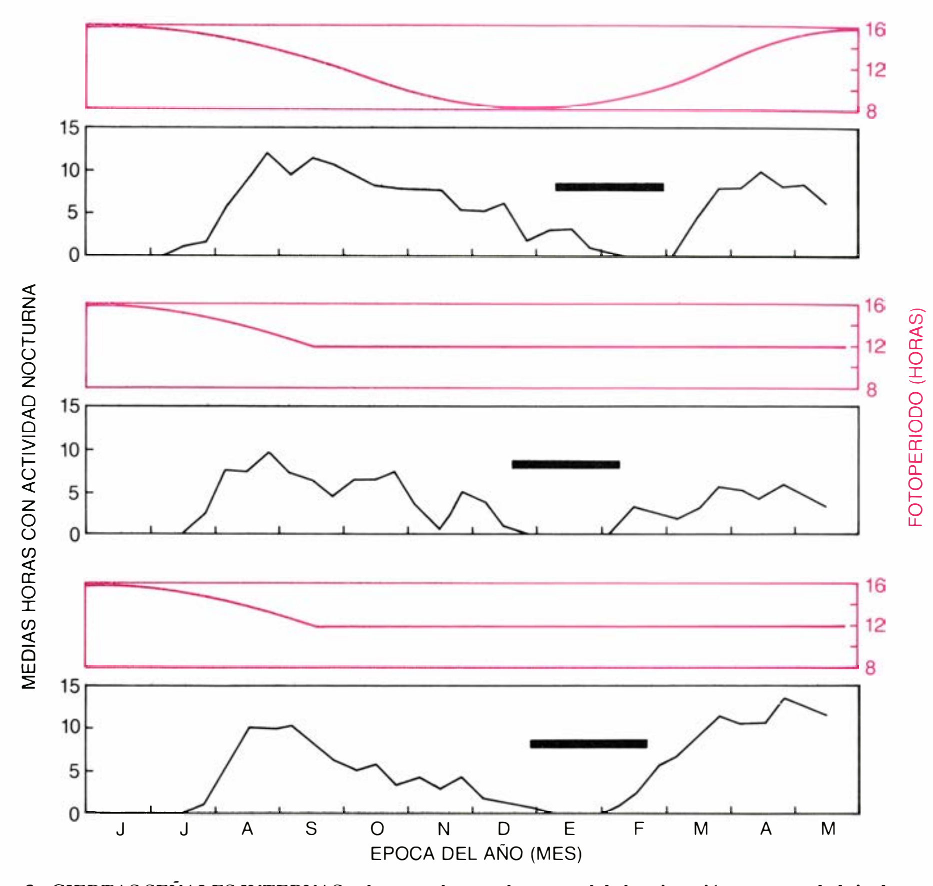


2. DOS ESPECIES DE CURRUCA figuran entre las aves estudiadas por el autor. A la izquierda, la curruca mosquitera (*Sylvia borin*); a la derecha, la curruca capirotada (*Sylvia atricapilla*). Ambas, aunque pequeñas y discretas son excelentes cantoras. El mapa que figura sobre cada ave muestra el área de reproducción (*gris*) y el área de invernada (*color*). En otoño, las aves dejan sus áreas de reproducción y ponen rumbo al Sur. La curruca mosquitera pasa el invierno en Centro y Suráfrica. La mayoría de las currucas capirotadas pasan el invierno en Europa meridional o en Africa, al norte del ecuador; algunas poblaciones orientales de esta misma especie emigran a Africa del Sur. En primavera, las currucas regresan a Europa y Asia septentrional para criar.

picales del hemisferio sur. Cría en Alemania y pasa el invierno en Africa ecuatorial.

Cuando emprendimos el estudio, se sabía que, para algunos migradores de pequeña escala, los cambios del fotoperíodo (el promedio del número de horas de luz por día) constituyen un importante estímulo del desencadenamiento de la migración. Por supuesto, en Africa ecuatorial, donde el mosquitero pasa el invierno, el fotoperíodo experimenta escasa variación estacional. Y puesto que fluctúa tan poco, parecía improbable que las señales de luz tuvieran gran cosa que ver con el inicio de la migración primaveral del ave. Tentaba plantear la hipótesis de que un reloj interno desencadenase el viaje de vuelta a casa sin la participación de estímulos externos.

Para verificar la hipótesis estudié el comportamiento de tres grupos de mosquiteros musicales. Los tres se encontraban en jaulas experimentales que permitían medir el grado de inquietud o desazón migratoria. En la mayoría de los mosquiteros la actividad nocturna suele manifestarse en forma de saltos: las aves no paran en las perchas y brincan por la jaula sin cesar. Bajo una de las perchas de la jaula experimental se instalaron microconmutadores. Cuando el ave saltaba a la percha, el microconmutador cerraba un circuito y un registrador cercano imprimía una marca sobre una cinta de papel. Contando el número de intervalos (medias horas, por ejemplo) en los que el ave se mostraba activa podía medirse la intensidad de la desazón.



3. CIERTAS SEÑALES INTERNAS subyacen al control temporal de la migración, como se dedujo de un experimento llevado a cabo con tres grupos de mosquiteros musicales (*Phylloscopus trochilus*). Cada par de gráficos ofrece los datos de un grupo de aves. La curva superior representa el fotoperíodo: el número de horas de exposición de las aves a la luz. La inferior, la intensidad de la inquietud migratoria (actividad nocturna desplegada por aves enjauladas cuando, de no estar presas, migrarían); la línea negra gruesa muestra la sincronización de la muda invernal, fija respecto de la migración. Se mantuvo un grupo de mosquiteros, durante todo el otoño y el invierno, en el laboratorio del autor, en Alemania, expuesto a la variación normal del fotoperíodo (*arriba*). A los otros dos grupos se les negó el estímulo estacional. Uno se llevó en septiembre al Zaire donde, por tratarse de un territorio próximo al ecuador, la variación del fotoperíodo es escasa (*medio*). El otro permaneció en el laboratorio, donde a partir de septiembre quedó expuesto a un fotoperíodo constante de 12 horas (*abajo*). Los tres grupos mudaban de plumaje y experimentaban la inquietud migratoria según el esquema anual de costumbre, señal de que esos acontecimientos los motiva un mecanismo de sincronización interno, y no señales estacionales, como la variación del fotoperíodo.

Mantuvimos en el laboratorio un grupo de aves, a temperatura constante y con un fotoperíodo, constante también, de 12 horas diarias. Pretendíamos comprobar de ese modo si las aves experimentaban, en ausencia de señales externas, el programa normal de acontecimientos estacionales asociados con la migración. A modo de control se llevó un segundo grupo al este del Zaire, a una latitud de unos dos grados Sur, dentro del territorio habitual de invernada del ave. Un tercer grupo pasó todo el invierno en nuestro instituto, en una habitación de grandes ventanales, expuesta, por consiguiente, a la variación del fotoperíodo típica de las regiones templadas. Mantuvimos los tres grupos en jaulas individuales para determinar el final de la inquietud migratoria de otoño y el comienzo de la primaveral.

Los resultados del experimento fueron diáfanos. En los tres grupos, la inquietud migratoria otoñal desapareció en diciembre y enero, y la primaveral empezó en febrero y marzo. Esa es la pauta típica de los mosquiteros musicales en la naturaleza. Es más, los tres grupos mudaron el plumaje en el momento apropiado. Las pautas de muda varían mucho entre especies. Sin embargo, en una misma especie la muda se realiza siempre en un momento fijo respecto del vuelo migratorio. Los mosquiteros musicales que viven en libertad sufren una muda completa a mediados de invierno, nada más llegar a los territorios de invernada; también mudaron entonces los tres grupos de aves enjauladas.

Que los mosquiteros, bajo condiciones constantes de luz y temperatura, siguieran el programa normal de inquietud y muda venía a sugerir que las aves recurren a señales internas para establecer el compás de sus actividades estacionales. Empero, no revelaba el mecanismo concreto de control temporal, pues los resultados del experimento podían interpretarse de dos formas. El control de los acontecimientos invernales podría responder a un mecanismo interno que se pusiera en acción cada primavera, en el territorio de cría. Semejante mecanismo se parecería a un reloj de arena de período anual; habrían de darle la vuelta los estímulos externos. Por el contrario, la programación de los acontecimientos estacionales reflejaría el funcionamiento de un mecanismo más parecido a un reloj eterno: un ritmo anual endógeno que operase durante muchos años sin necesidad de darle cuerda.

El experimento de 1966 no nos permitió decidir qué conclusión era la correcta, pero mi colega Peter Berthold y yo ideamos otros que sí nos facultó. El ave estudiada en los nuevos experimentos fue la curruca mosquitera (*Sylvia borin*), especie cuyas pautas de migración se parecen mucho a las del mosquitero musical. Expusimos tres grupos

de currucas mosquiteras a un fotoperíodo constante durante tres años. En uno de ellos la duración de la luz era de 10 horas; en el segundo de 12 y, en el tercero, de 16. Los resultados fueron idénticos en todos los casos: la inquietud migratoria, la muda del plumaje y otros comportamientos relacionados siguieron una pauta anual regular durante todo el experimento.

Posteriormente realizamos un ensayo, más largo aún, con currucas mosquiteras y currucas capirotadas (*Sylvia atricapilla*). Como antes, expusimos las aves a un período invariable de luz diaria. En esta ocasión, sin embargo, atendimos sólo a la muda del plumaje, que ha resultado ser uno de los indicadores más fiables del compás cíclico, anual. En la naturaleza, las currucas capirotadas y las mosquiteras cambian de pluma dos veces al año. En condiciones experimentales, durante ocho o incluso más años, ambas especies mantuvieron esa pauta de muda. Los resultados de las dos series de experimentos indicaban sin ambigüedad que las aves en cuestión disponían de un cronómetro interno, rítmico, que hacía presentarse los acontecimientos estacionales en el orden correcto.

Resulta notable que, en condiciones ambientales constantes, el período del ritmo interno no fuera exactamente de un año. Por ejemplo, las mudas de invierno de las currucas capirotadas se producían más o menos cada diez meses, no cada 12. Tal es la razón de que se califiquen esas oscilaciones endógenas de circanuales, término que, derivado del latín, viene a significar "alrededor de un año". Hace 25, Eric T. Pengelley y Kenneth C. Fisher, de la Universidad de Toronto, describían por vez primera los ritmos circanuales en los citelos [véase "Annual Biological Clocks", por Eric T. Pengelley y Sally J. Asmundson; SCIENTIFIC AMERICAN, abril de 1971]. Desde entonces se han descrito fenómenos periódicos similares en una amplia gama de organismos, incluidos los celentéreos, insectos, peces, anfibios, reptiles y mamíferos, además de aves.

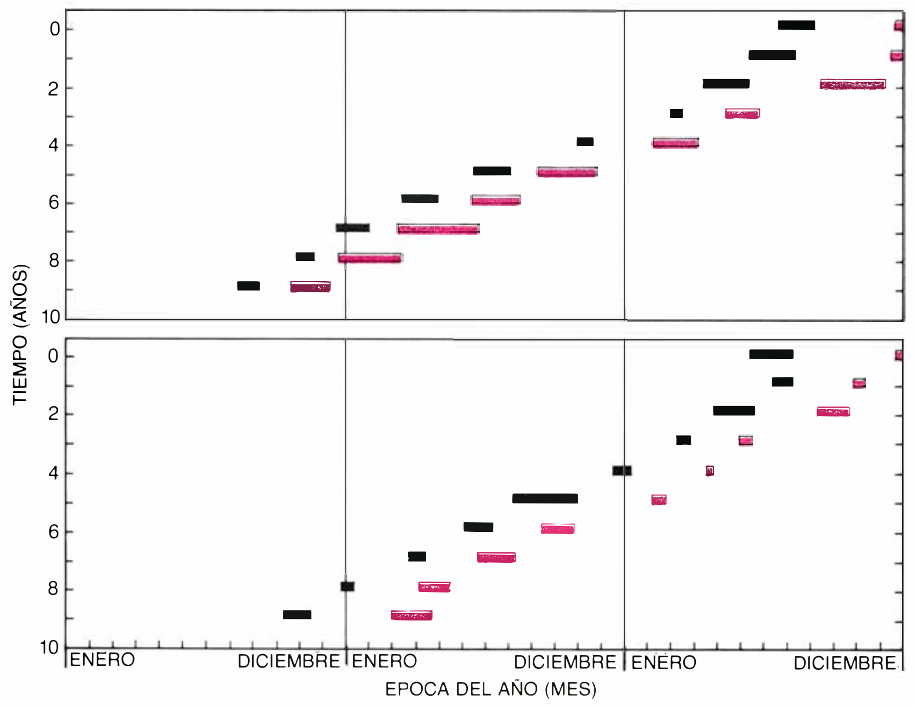
De esas oscilaciones observadas por nosotros en la curruca mosquitera y la curruca capirotada y su desviación del año cabía extraer dos consecuencias importantes. Primero, demostraban que los ritmos circanuales no podían deberse a la influencia de estímulos estacionales, no controlados, generados en el laboratorio. Si los hubiera provocado un estímulo de esa índole, habrían tenido un período de un año; que se desviaran del año indicaba que eran internos y espontáneos.

La segunda consecuencia es que si las actividades estacionales, como la muda de la pluma y la migración, se realizan siempre en la misma época del año, debe existir algún factor que ponga los ritmos internos circanuales en concordancia con el año solar. No costó dar con el sincronizador. Logramos que un grupo de currucas experimentara la desazón migratoria y la muda del plumaje según un ciclo anual exacto simulando la variación anual del fotoperíodo a 40 grados de latitud norte, que corresponde al territorio de cría europeo de la curruca. Más sorprendente aún fue que pudiera comprimirse el ritmo circanual a la mitad de su duración normal aumentando y disminuyendo el fotoperíodo según un ciclo de seis meses, en vez de doce.

Habíamos dejado sentado, por tanto, que el ritmo circanual, oportunamente modificado por factores ambientales, marca el compás de la pauta estacional de acontecimientos. Parecía importante averiguar si los ritmos endógenos incidían también en las características del vuelo migratorio, como la duración e incluso la meta final. Para estudiar esas cuestiones debíamos emplear aves de las que se supusiera que no habían aprendido, por experiencia, una ruta migratoria específica.

Casi todas las aves de nuestros experimentos eran criadas por seres humanos. Cuando cumplen la semana de edad, se las separa de los padres y pasan al cuidado del personal del laboratorio; éste las alimenta con gusanos, larvas de insecto y otros apetitosos bocados. Tal proceso, que podemos llamar de cría a mano, ofrece por lo menos dos ventajas apreciables en un trabajo como el nuestro. Primera: cuando las aves se colocan en una instalación experimental, están ya acostumbradas a la jaula; no muestran el errático comportamiento observado en las aves salvajes enjauladas por vez primera. Dicho comportamiento interferiría los resultados experimentales, sobre todo los que se centran en el estudio de la inquietud migratoria.

La otra gran ventaja es que, si se crían las aves a mano, se conoce su edad, procedencia e historia. De este modo, cuando abordamos la influencia de los ritmos circanuales en los pormenores de la pauta migratoria, pudimos excluir la influencia del aprendizaje: tratábamos con aves de las que sabíamos que nunca habían emigrado.



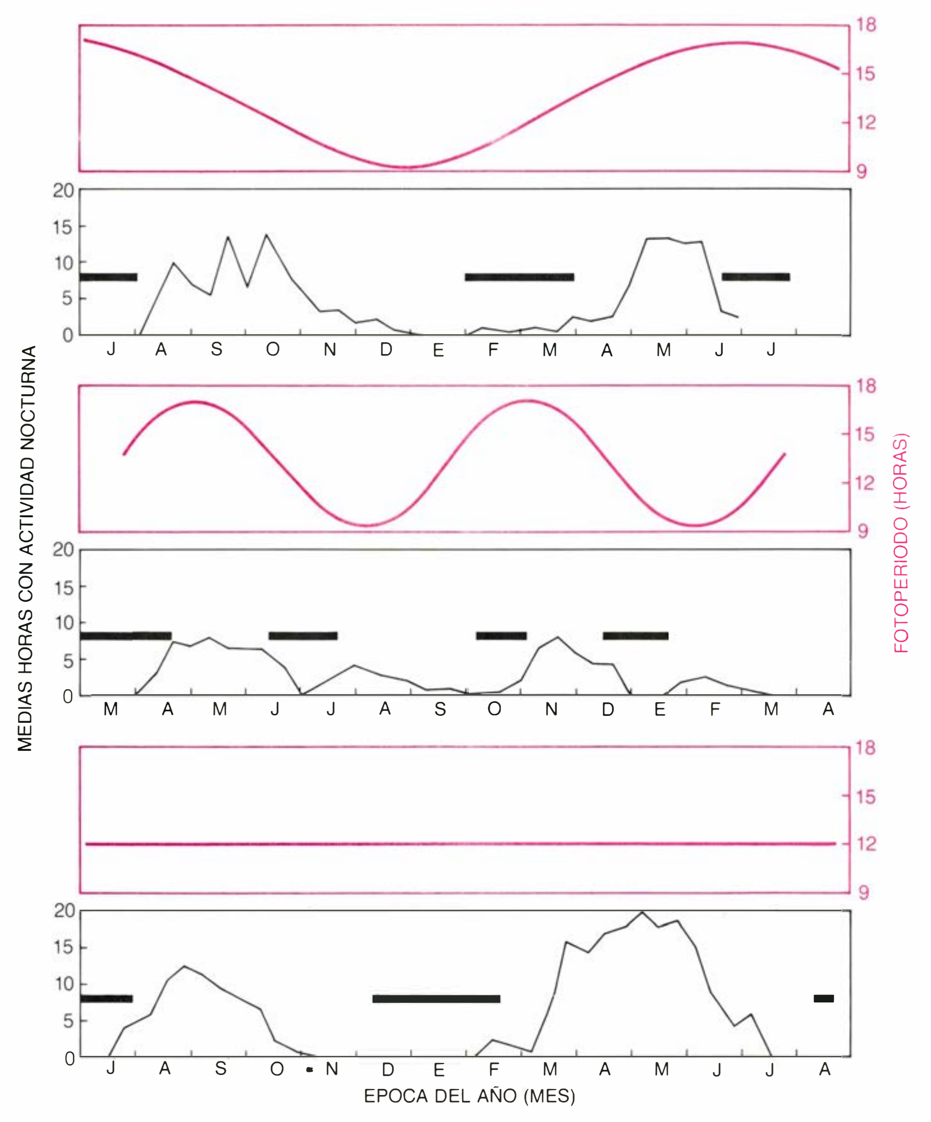
4. RITMOS CIRCANUALES controlan el compás de los acontecimientos estacionales en algunas especies de aves, como lo demostró un prolongado experimento con currucas capirotadas y currucas mosquiteras. Ambas especies cambian de pluma dos veces al año, una en invierno y otra en verano. Se mantuvieron las aves en condiciones de temperatura y fotoperíodo constantes, al tiempo que se observaba la sincronización de las mudas. Los resultados para una curruca capirotada figuran en el gráfico superior; los de una curruca mosquitera, en el inferior. Las barras horizontales negras corresponden a las mudas estivales; las de color, a las de invierno. Las mudas se presentaron regularmente durante 10 años bajo condiciones ambientales constantes. Sin embargo, el período entre los acontecimientos anuales era sólo de unos 10 meses. (Por ejemplo, las mudas invernales se distanciaban 10 meses.) Así pues, las aves mudaban algo más temprano cada año. Dado que el período de esos ritmos internos no es exactamente de un año, se les denomina circanuales.

Según descubrimos, las aves inexpertas presentaban pautas de inquietud que reflejaban el curso del vuelo migratorio de las silvestres. Los mosquiteros musicales enjaulados mostraban el grado máximo de desazón en agosto y septiembre, justo cuando las aves de vida libre cruzan el Mediterráneo y el Sáhara. Después, la actividad nocturna de las aves enjauladas disminuía, igual que se modera el vuelo migratorio cuando las aves salvajes se aproximan a sus cuarteles de invierno. La inquietud desaparecía por la época en que los mosquiteros en libertad alcanzan el territorio de invernada.

Tales resultados daban a entender que el mosquitero musical no "aprendía" mediante el curso temporal completo de su migración. Trabajos ulteriores permitieron generalizar esa conclusión. Descubrimos que, en las aves enjauladas inexpertas, la duración de la inquietud migratoria era proporcional a la distancia cubierta por la especie en su vuelo migratorio otoñal. Mis colegas Peter Berthold y Ulrich Querner demostraron que igual correlación se daba en el seno de una misma especie. Compararon currucas capirotadas que criaban en Finlandia, Alemania, Francia y las islas Canarias. Las aves finesas, que deben realizar el viaje más largo para llegar hasta Africa central, presentaban la actividad máxima; seguíanles alemanas, francesas y canarias, en este orden, que viajan a distancias progresivamente más cortas.

Berthold y Querner explotaron la variabilidad entre las subpoblaciones de curruca capirotada para demostrar que el curso temporal de la migración está programado genéticamente. Su estrategia consistió en cruzar las currucas capirotadas de Alemania con las procedentes de las islas Canarias para ver qué pauta cronológica de desazón migratoria exhibían los descendientes. Si, como suponían, la pauta respondía a dictados genéticos, entonces los descendientes presentarían un grado de inquietud intermedio con respecto a los dos grupos parentales. La razón estriba en que todo comportamiento de la complejidad de la migración probablemente sea producto de la actividad de varios genes, y la herencia de los caracteres poligénicos adopta a menudo la forma de un promedio de los extremos parentales.

La cría de currucas capirotadas en cautividad es una tarea ardua y frustrante, pero Berthold y Querner se empecinaron en ello hasta que obtuvieron 32 híbridos de padres alemanes y canarios. Criaron las aves a mano y las mantuvieron en jaulas hasta su primer otoño, momento en que se midió la intensidad de su inquietud migratoria. Los resultados fueron los pronosticados por la hipótesis poligénica: los híbridos presentaron una pauta de actividad nocturna casi intermedia entre las correspondientes a las poblaciones parentales.

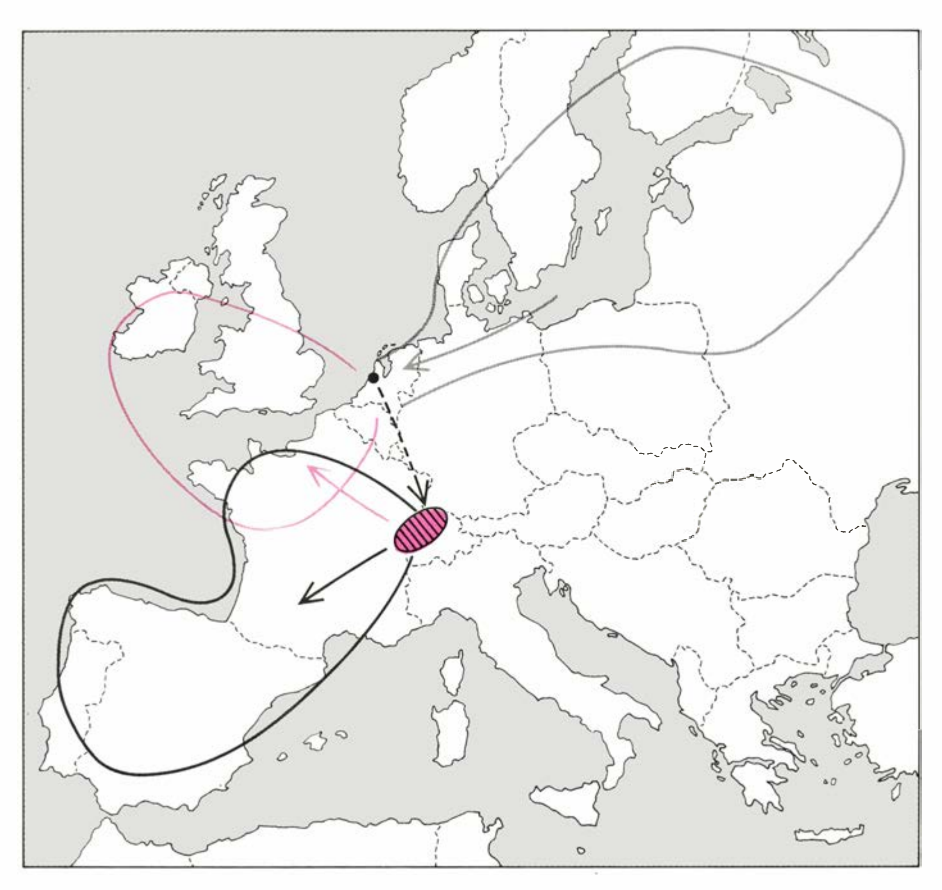


5. SINCRONIZACION DE LOS RITMOS circanuales por medio de estímulos externos. Dado que el período de los ritmos circanuales no es de un año justo, deben sincronizarse de modo que los acontecimientos fisiológicos que subyacen a la migración coincidan con las estaciones. El autor demostró que los ritmos endógenos pueden modificarse por la variación del fotoperíodo. Expuso un grupo de currucas a los cambios estacionales del fotoperíodo: experimentaban inquietud migratoria y mudaban de plumaje según el esquema anual acostumbrado (*arriba*). Mantuvo un segundo grupo bajo condiciones constantes; exhibían un ritmo circanual con un período de más de doce meses (*abajo*). Expuso un tercer grupo a un fotoperíodo que variaba según un ciclo de seis meses: ritmo de muda e inquietud migratoria se comprimieron a seis meses (*centro*).

El experimento de cruzamiento llevado a cabo por Berthold y Querner demostraba, sin lugar a dudas, que el curso temporal de la inquietud migratoria estaba determinado por un programa genético. Pronto se evidenció que ese programa temporal ayudaría a las aves jóvenes a localizar el territorio de invernada, lo cual resulta de importancia obvia para la supervivencia.

En las especies que viajan en grupo, los migradores inexpertos pueden seguir a los guías más avezados. Sin embargo, muchas emigran en solitario, incluidas la mayoría de las aves con que trabajábamos. En esas especies, cada ave ha de ser capaz de encontrar el territorio de invernada por su cuenta y riesgo, haya o no hecho el viaje con anterioridad. Está claro que, si un ave no ha emigrado antes, el aprendizaje no puede ayudarle a dar con el camino; por tanto, algún otro mecanismo ha de servir para navegar. Se van acumulando pruebas de que tal mecanismo es un programa temporal.

Es cosa sabida desde hace algún tiempo que los novatos se valen de métodos de navegación diferentes de los empleados por los migradores más via-



6. **SISTEMA DE NAVEGACION utilizado por las aves inexpertas, que difiere del que emplean los adultos que han efectuado el vuelo migratorio al menos una vez. A. C. Perdeck, del Instituto holandés de Investigación Ecológica, demostró esa discrepancia. Perdeck capturó estorninos a medio camino de su ruta migratoria de otoño (*gris*), desde los territorios de cría, en el Báltico, hasta las áreas de invernada, en Francia e Inglaterra. Transportó las aves a Suiza y las soltó. Los migradores expertos (*color*) compensaron el desplazamiento y volaron hacia el área normal de invernada. Las aves sin experiencia (*negro*) siguieron volando en la dirección original, con un rumbo que las llevó a la península Ibérica. Al parecer, las aves inexpertas no están preparadas para una verdadera navegación; tan sólo pueden volar siguiendo en una dirección fija de la brújula.**

jados. Así lo demostraron, con extraordinaria claridad, los trabajos de A. C. Perdeck, del Instituto holandés de Investigación Ecológica. Estudió el estornino pinto, que medra en las costas del mar Báltico de Suecia, Finlandia, Letonia, Lituania, Polonia, Alemania y Dinamarca. En uno de sus experimentos, Perdeck capturó 11.000 estorninos cerca de La Haya y los transportó en avión a Suiza, donde se anillaron y soltaron.

Se recapturaron más tarde 300 aves; al situar en el mapa los puntos de recuperación se advirtió una distribución sorprendente. Las aves que habían emigrado con anterioridad por lo menos una vez sabían compensar el desplazamiento hacia el sur. Aparecían dispersas a lo largo de una trayectoria de vuelo que conducía a los territorios de invernada, trayectoria que suponía dar un giro brusco hacia el norte desde el rumbo original. Por el contrario, las aves jóvenes se localizaban a lo largo de una trayectoria de vuelo más o menos paralela a la normal, que las llevaba excesivamente al sur de la meta habitual [*véase la figura 6*]. Había que concluir que las aves sin experiencia eran incapaces de compensar el desplazamiento y se limitaban a volar en la dirección original después de soltarlas.

El resultado de Perdeck indica que las aves jóvenes, cuando emigran, sólo pueden volar en una dirección general. Se considera primitivo el sistema de navegación que sólo acomode un programa de vuelo, pero ello no significa que carezca por completo de valor práctico. Si el programa inicial es correcto y las perturbaciones externas (como las producidas por el viento) no son demasiado intensas, todo lo que el aviador necesita para alcanzar la meta es un reloj que le indique cuándo empezar a volar y cuándo detenerse.

Diversos experimentos de campo han proporcionado resultados conformes con la hipótesis de que las aves inexpertas dependen de un reloj de navegación. La demostración más clara fue quizá la realizada por W. Rüppell y E. Schüz, del Vogelwarte Rossitten, precursor de nuestro instituto. Capturaron cornejas negras jóvenes hacia la mitad de su viaje migratorio desde los territorios de cría, en el Báltico, hasta las áreas de invernada de Alemania septentrional. Anillaron las aves y las sacaron del área de distribución normal de invierno, soltándolas luego. Al recapturar, durante el invierno, algunas cornejas anilladas, se descubrió que habían seguido emigrando con el programa inicial, que ahora las llevó fuera del territorio de invernada. Además, habían volado una distancia poco más o menos igual a la que las separaba de la meta. Diríase que el reloj migratorio siguió marcando el tictac después del desplazamiento; las cornejas jóvenes volaron la distancia justa, pero lejos de la meta.

Los resultados de Perdeck, combinados con los de Rüppell y Schüz, ayudaron a aclarar la diferencia que distingue los sistemas de navegación de las aves con y sin experiencia. Las que ya han emigrado disponen de capacidad para una verdadera navegación: desplazadas, pueden corregir el rumbo y dar con la meta. (Cómo lo logran es algo que todavía no sabemos.) Las aves inexpertas no tienen tal capacidad; al parecer, cuentan con un solo programa y con un reloj de vuelo presincronizado. Hasta hace poco se desconocía la naturaleza del cronómetro de navegación. Empero, nuestro trabajo con currucas y mosquiteros indica que el programa genético, que especifica el curso temporal de la inquietud migratoria, posee ese reloj. El ritmo fisiológico que subyace a la inquietud observada en las aves enjauladas le dicta al ave cuándo ponerse en camino, cuándo ejercer el máximo esfuerzo, cuándo frenar y en qué momento debe cesar el vuelo.

Aunque los sistemas de navegación difieren considerablemente, novatos y expertos comparten la necesidad de conocer el programa de vuelo correcto. Lo curioso es que, por lo que se ve, los ritmos circanuales también pueden ayudar a la resolución de ese problema. Durante algún tiempo, los resultados experimentales disponibles señalaban que la elección de un programa brújula por parte del migrador venía determinada más por su estado fisiológico que por señales externas, así la información geográfica o la posición de las estrellas. Entre esos datos figuran algunos obtenidos por Stephen T. Emlen, de la Universidad de Cornell, [véase "The Stellar-Orientation System of a Migratory

Bird", por Stephen T. Emlen; SCIENTIFIC AMERICAN, agosto de 1975].

Emlen expuso un grupo de *Passerina cyanea* a dos ciclos fotoperiódicos de un año. Los cambios de fotoperíodo, en el sentido de acelerarlo, adelantaban el compás de la muda y de la inquietud migratoria en las aves experimentales. Con el tiempo, éstas llegaron a un estado fisiológico "otoñal" en primavera. Se colocaron entonces en jaulas de orientación en un planetario con un cielo de primavera simulado. A pesar de las señales externas, las aves respondían a su estado interno y, consecuentemente, saltaban hacia el sur, como si intentaran iniciar el vuelo migratorio de otoño.

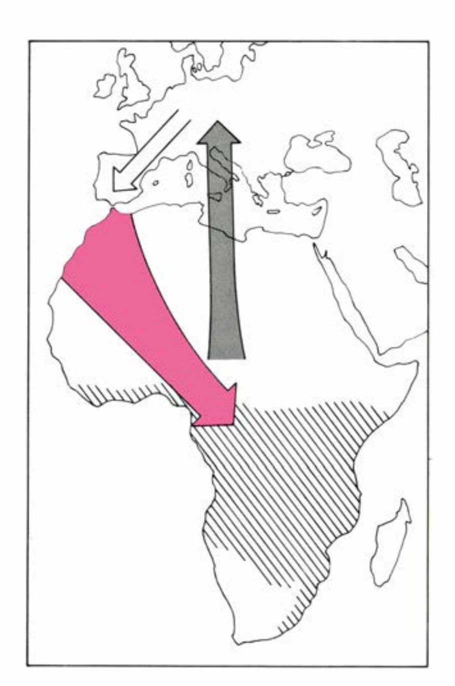
Desde la perspectiva de los trabajos que venimos realizando mis colegas y yo, los resultados de Emlen pueden interpretarse de dos formas. La tría de la dirección incorrecta en los *Paserina cyanea* podría deberse al efecto del fotoperíodo, o al del ritmo circanual subyacente. Necesitábamos saber cuál era el factor responsable. Así, en colaboración con Wolfgang Wiltschko, de la Universidad de Frankfurt, diseñé un protocolo de experimentos que desenmarañaran los dos efectos. Realizamos las pruebas con 59 currucas mosquiteras criadas a mano. Figuran esos pájaros en un amplio grupo de aves europeas que comparten una ruta migratoria básica. Abandonan el territorio de cría en otoño con rumbo Suroeste. Sin embargo, llegadas a la península Ibérica o a Africa del norte, doblan al Sur o Sur-Sureste, y mantienen ese rumbo hasta alcanzar el territorio invernal. El vuelo de regreso en primavera sigue una ruta más directa: las aves vuelan en línea recta de Sur a Norte.

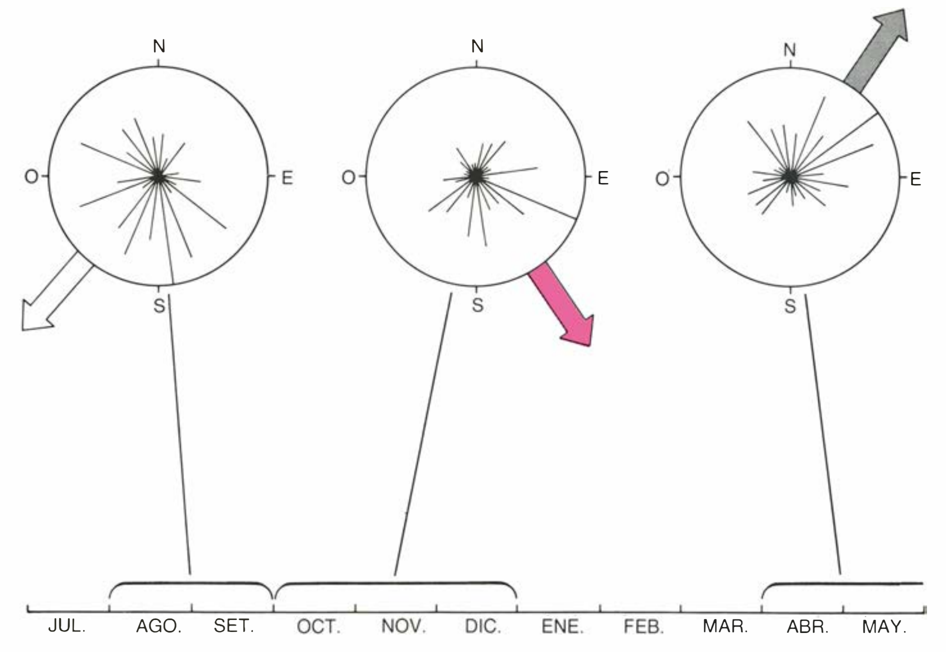
Como en nuestro trabajo anterior, para aislar el efecto del ritmo circanual necesitábamos minimizar las fluctuaciones estacionales experimentadas por las aves. Manteníamos las currucas en cámaras con temperatura constante y un fotoperíodo invariable.

Para poner a prueba la tría de dirección en un ave, ésta ha de poder orientarse en relación a los puntos de la brújula. Wiltschko había demostrado previamente que la curruca mosquitera podía orientarse por medio de las estrellas o del campo magnético terrestre. El mapa celeste cambia con la estación. Puesto que nos proponíamos eliminar todas las señales estacionales, a las aves no se les permitió ver el cielo. Por otro lado, el campo magnético de la Tierra experimenta una variación estacional muy pequeña. No tomamos medida alguna para aislar las aves del campo magnético y, por consiguiente, pudieron orientarse con respecto al mismo.

El montaje experimental estaba casi completo. Se habían aislado las currucas de las señales estacionarias y se les había proporcionado los medios para percibir las direcciones de la brújula. Ahora bien, necesitábamos un método para comprobar las preferencias de dirección de las aves enjauladas durante el período migratorio. A ese fin, durante el año, y a intervalos regulares, colocábamos las currucas en jaulas especiales, octógonos de un diámetro aproximado de un metro. Ocho perchas radiaban del centro. Unos microconmutadores montados bajo las perchas nos permitían averiguar cuáles preferían las aves durante la noche, en cada fase del ciclo migratorio.

Los resultados fueron absolutamente esclarecedores. En agosto y septiembre, cuando las currucas mosquiteras salvajes se dirigen al suroeste en lo que constituye la primera bordada de su viaje migratorio triangular, las aves enjauladas tendían a saltar a las perchas de la esquina suroeste de la jaula. Durante la segunda mitad del período migratorio de otoño, cuando las aves en migración doblan hacia el Sur y apun-





7. PREFERENCIAS DE DIRECCION en las currucas mosquiteras. Se encuentran bajo el control de ritmos circanuales, como quedó demostrado en un experimento realizado por el autor y Wolfgang Wiltschko, de la Universidad de Frankfurt. La curruca mosquitera emigra siguiendo una trayectoria triangular (*mapa izquierdo*). En septiembre deja Europa y pone rumbo al suroeste. A finales de otoño cambia de dirección a la altura de la península Ibérica y se dirige al sureste, hasta Africa ecuatorial. El viaje de regreso, en primavera, sigue una ruta directa Sur-Norte. Los tres círculos de la derecha muestran las preferencias de dirección en las currucas mosquiteras mantenidas bajo condiciones constantes de luz y temperatura, en jaulas de orientación como la ilustrada en la figura 1. Cada círculo corresponde a los resultados obtenidos en una época del año. Las líneas que parten de los centros indican la frecuencia con que las currucas prefieren determinadas direcciones de la brújula. Cada línea representa 15 grados de la brújula. La dirección escogida más a menudo figura como radio del círculo, mostrándose las demás en proporción a ella. La flecha grande que se proyecta desde el borde del círculo señala la dirección general que resulta de sumar las elecciones de las aves enjauladas. En agosto y septiembre, las aves preferían la dirección Suroeste; en octubre, noviembre y diciembre, la Sureste y, en abril, mayo y junio, la Noreste. Los resultados vienen a coincidir con las direcciones migratorias que siguen las aves libres en esas épocas del año.

tan a Africa central, las enjauladas preferían posarse en las partes sur y sureste de la jaula. En primavera, cuando los migradores avanzan en línea recta hacia Europa desde sus cuarteles ecuatoriales de invierno, las aves enjauladas escogían muy a menudo las perchas septentrionales.

Parece evidente que los ritmos circanuales pueden controlar la dirección de la migración, así como su inicio y su estructura temporal. En realidad, el resultado global del trabajo realizado en mi laboratorio durante las últimas dos décadas hace hincapié en los efectos del reloj interno, espontáneo. Con todo, no se vaya a suponer que el funcionamiento de los ritmos circanuales elimina el influjo de las señales externas. Como he descrito anteriormente, los cambios de fotoperíodo o, en aves adultas, el aprendizaje de la navegación, pueden modificar los ritmos circanuales. Algunos trabajos recientes se han propuesto desentrañar las modificaciones del programa interno: mi colega Herbert Biebach ha demostrado que las modificaciones pueden ser bastante específicas.

Biebach se centró en los cambios de peso que experimentan las aves mientras emigran. Antes de partir, los migradores suelen comer más que de costumbre, para dotarse de reservas de grasa que utilizar como combustible durante el viaje. La ganancia de peso puede ser considerable. En el campo, las currucas mosquiteras llegan a engordar desde un peso inicial de 16 gramos hasta los 30 gramos y, en el laboratorio (donde tienen a mano una provisión ilimitada de alimento), la ganancia puede ser aún mayor.

Biebach capturó canoros en dos lugares del Sáhara. En un caso, se trataba de una zona artificial de descanso hecha con unos cuantos arbustos plantados en la arena. Las aves, atraídas por los arbustos, bajaban a ellos, pero no había ni agua ni comida. El otro punto era un oasis con abundante vegetación y una rica vida insectícola. En ambos sitios Biebach capturó muchas de las aves que acudían por la mañana, las pesó, las anilló y las soltó luego. Algunas se recapturaron después para volverlas a pesar. Resultó que las aves se comportaban de forma distinta en uno y otro lugar. En la trampa del desierto, donde no había alimento, todas las aves partían al atardecer del día de la llegada. En el oasis, sin embargo, una proporción bastante grande de las aves llegadas se quedaba durante más de un día. Los datos de peso mostraban que las aves que se quedaban tendían a ser aquellas cuya reserva de grasa a la llegada se encontraba más agotada.

Los hallazgos de Biebach indicaban que, cuando al descenso de las reservas de grasa se le sumaba la oportunidad de reponerlas, se inhibía el programa migratorio endógeno que, de lo contrario, impulsaría a las aves a continuar el vuelo. Tal conclusión se ha corroborado en el laboratorio. Los estudios experimentales demuestran que ni el agotamiento de las reservas de grasa, ni la oportunidad de alimentarse, son de por sí suficientes para interrumpir el funcionamiento del programa; han de coincidir ambos factores para que se inhiba el impulso migrador. En cuanto el ave alcanza cierto peso corporal mínimo, el impulso reaparece.

Que el programa endógeno quede relegado a un segundo término por la acción de estímulos externos ofrece ventajas manifiestas. Si no cupiera modificación alguna de los ritmos circanuales, el ave cuyas reservas de grasa se hubieran agotado podría morir de hambre, empujada por un implacable impulso migrador. Si éste admite una atenuación temporal, el ave en migración reabastecerá sus reservas de grasa antes de reemprender el vuelo. En cierto sentido, el comportamiento que tan sutilmente ha desvelado Biebach es similar al sistema de navegación de las aves experimentadas. En ambos casos, un mecanismo bastante inflexible resulta desplazado por otro sistema que puede ajustar el programa interno a condiciones externas, específicas.

Una de las principales tareas que afrontan los estudiosos de la migración es averiguar precisamente cómo interactúa el programa endógeno con los estímulos externos para producir el comportamiento migratorio observado. Otro problema de igual trascendencia concierne a la identificación de la base fisiológica de los ritmos circanuales. Mis colegas y yo trabajamos en ambas cuestiones, pero nos puede llevar muchos años hallarles respuesta. Después de todo, ha costado dos décadas poner en claro que los ritmos endógenos constituyen el armazón general del comportamiento migratorio. Los detalles del ciclo migratorio son susceptibles de diversas modificaciones para acomodarse a las circunstancias externas, pero el impulso que mueve a la gran V de barnaclas del Canadá, así como al atropellado grupo de estorninos, procede de ritmos anuales generados en lo más profundo del organismo.

Nuevas estrategias para la energía nuclear

Una estrategia para liberar la energía nuclear del punto muerto actual podría surgir de la nueva generación de reactores de baja potencia, de probada seguridad

Richard K. Lester

Los reactores nucleares generan alrededor del 15 por ciento de la producción eléctrica mundial. Francia, Alemania Federal, Japón y otras naciones que dependen, en buena parte, de la energía nuclear tienen en proyecto un nuevo impulso del sector. La misma situación se presenta en la mayoría de naciones de Europa oriental y en otros países menos desarrollados. En Estados Unidos, donde se desarrollaron los reactores de agua ligera (LWR) que suministran la mayor parte de la electricidad mundial, y donde, con mucho, se ha construido el mayor número de centrales nucleares de dicho tipo, la energía nuclear se encuentra en una fase de estancamiento. Las compañías eléctricas que, en los primeros años setenta, se embarcaron en un vasto programa de proyectos nucleares, han abandonado, una década más tarde, la idea de construir una central nuclear más. A pesar de la prevista necesidad de una mayor potencia instalada para finales de siglo, la energía nuclear no puede reclamar hoy "ser la opción a utilizar en última instancia", por usar una frase del ex-presidente Carter. Para las compañías eléctricas ésa no es, en absoluto, una opción.

El éxito de los proyectos nucleares, dentro y fuera de los Estados Unidos, ha confirmado que no existe nada que, de suyo, resulte inviable en este contexto. Además, aunque se ofrezcan otras fuentes alternativas de energía para las compañías eléctricas y estén en curso de desarrollo técnicas distintas de la nuclear, el abandono de ésta conllevaría una producción eléctrica menos eficaz y menos benigna desde el punto de vista ambiental. Abunda todavía el petróleo y el gas, pero su crecimiento será más lento y costoso. No es fácil una combustión limpia del carbón: la lluvia ácida y otros subproductos ocasionan graves daños a la biosfera no siempre controlables. Además, el dióxido de carbono que se libera en la combustión de los combustibles fósiles puede producir serios problemas en el clima. En otro sentido, está por demostrar que la técnica solar para la producción de energía eléctrica le convierta en económicamente competitiva, salvo en casos muy aislados y para usos concretos.

Ante tal situación, el abandono de la energía nuclear no sería una decisión acertada. Ahora bien, se halla inmersa en tantos y tan distintos problemas, por lo que a Estados Unidos se refiere, que, para revitalizarla, habría que acometer profundos cambios en la organización industrial y en la propia tecnología nuclear. Transformaciones, además, nada fáciles. Pero esas oscuras perspectivas y esos problemas de la industria nuclear han creado un estado de cosas en que se hace imperioso reexaminar planteamientos anteriores y explorar nuevas iniciativas. ¿Qué pasos habrá que dar para que cambie la suerte de la energía nuclear? ¿Qué probabilidad existe de que gobierno e industria se decidan a darlos?

Se acepta que la fusión parcial del reactor nuclear de Three Mile Island, en 1979, cortó el camino de éxitos de la industria nuclear. No se detectaron secuelas en la salud pública, pero el hecho de que se produjera el accidente, junto a frecuentes acciones carentes de tacto por parte de las autoridades después del mismo, asestaron un duro golpe a la confianza pública en la seguridad de esta tecnología. Además, al quedar afectada la viabilidad financiera de la compañía propietaria de la central, el accidente redobló la sensibilidad de las compañías eléctricas ante los riesgos de nuevas inversiones en el campo nuclear.

La verdad es que razones de mayor peso habían ido apagando el entusiasmo de las empresas antes del accidente. Las centrales que se construyen hoy, con una potencia de 1000 megawatt, son grandes y costosas. El sistema de control económico a que las compañías están sujetas ha servido para disuadirlas de acometer grandes inversiones de capital en centrales de cualquier clase. Las comisiones de las empresas de utilidad pública y la Comisión Federal Reguladora de la Energía ejercen, mediante un mecanismo de regulación de tarifas, el control de las amortizaciones de inversión de las compañías privadas, las cuales suministran el 80 por ciento de la electricidad en los Estados Unidos. Las entidades reguladoras garantizan a las empresas franquicias monopolistas, pero les obligan a ofrecer a todos los consumidores un adecuado suministro eléctrico a un coste razonable; en contrapartida, están autorizadas a establecer unos precios que den unos beneficios razonables a sus accionistas.

Durante gran parte del siglo, en que el precio de la electricidad fue decreciendo, el acuerdo funcionó bastante bien. Desde principios de los años setenta, sin embargo, en que el coste de la energía procedente de las nuevas centrales comenzó a superar el histórico precio medio de la electricidad, la presión de los consumidores aumentó y las relaciones entre compañías y entidades reguladoras se agriaron. Las empresas afirman que los ingresos no han aumentado al mismo ritmo que los costes. Un número creciente de comisiones estatales ha venido impidiendo que las compañías recuperen, mediante au-

mento de sus tarifas, los costes de construcción para proyectos que se consideraron, según la opinión pública, innecesarios o mal dirigidos (muchos de los proyectos en cuestión se referían a centrales). Así se ha ido erosionando gravemente la confianza del inversor y se ha llegado a que la mayoría de las compañías descartasen nuevas construcciones de cualquier tipo.

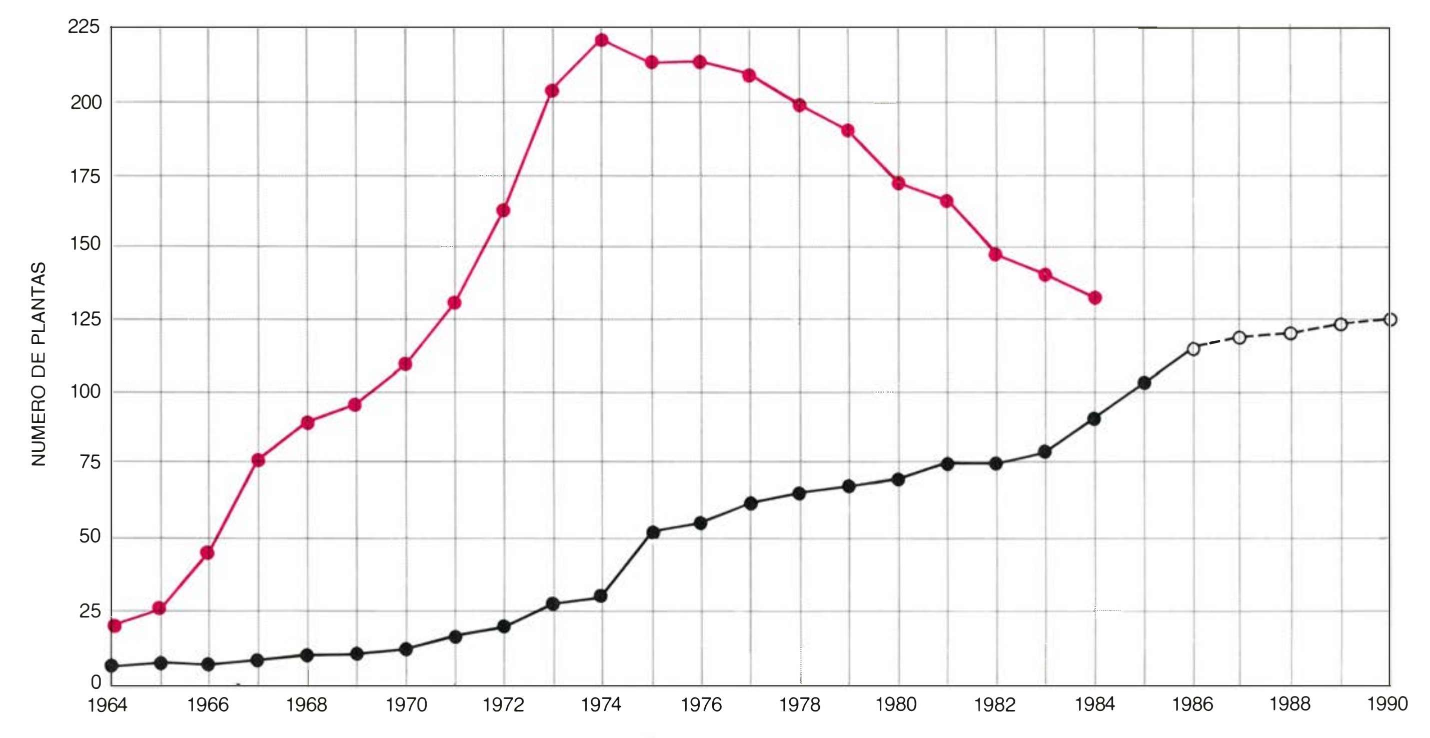
De los factores responsables de los infortunios de este sector industrial, muchos dimanan de la naturaleza de la energía nuclear. Las dificultades que entraña el almacenamiento de los residuos radiactivos desaconsejan nuevos encargos de reactores. Además, continúa paralizado el programa previsto para la ubicación de los residuos de alta y baja actividad. Añádase que, salvo algunas notables excepciones, el rendimiento de las centrales ha ido por debajo del esperado. El factor medio anual de capacidad (porcentaje de la potencia total teórica producido en el transcurso de un año) no ha superado, durante varios años, el 60 por ciento, cifra bastante menor que la conseguida por otros países: Japón, Suecia, Suiza y Alemania Federal.

La razón más importante que aducen las compañías para considerar poco atractiva la energía nuclear reside en el elevado coste de una central. Las 22 centrales nucleares que entraron en servicio en el período comprendido entre 1983 y 1985 ponen de manifiesto el problema; los costes de construcción alcanzaron los 2300 dólares por kilowatt instalado. Entre tanto, las centrales de carbón que entraron en servicio durante un período similar costaron, de media, menos de 1000 dólares por kilowatt. Tamaña diferencia de capital entre centrales nucleares y de carbón compensa, con mucho, los costes del combustible nuclear, menores, y se ha invertido su tradicional ventaja sobre el carbón.

Hay excepciones. Varias centrales recientes son ya competitivas con las de carbón. Pero es la norma, y no la excepción, lo que conforma las esperanzas de las compañías; de ahí que el panorama de construcción de nuevas centrales se está volviendo cada vez más desalentador. Otras 10 centrales entrarán en servicio a lo largo de este año y, según estimaciones recientes, su coste medio excederá los 3100 dólares por kilowatt.

¿Por qué se han encarecido tanto las centrales nucleares? Entre las causas hay que citar la dilatación del tiempo necesario para su construcción. Las centrales que entraron en servicio entre 1983 y 1985 tardaron una media de 10 años en techarse, frente a los cinco de las centrales de carbón que lo hicieron en el mismo período y los seis de proyectos nucleares similares de Francia y Japón. Los tipos de interés y la inflación de los últimos años han sido determinantes. Considérese el caso de la construcción simultánea de dos centrales nucleares en unas condiciones económicas más benignas, como las actuales. Si una se termina en cinco años y la otra necesita 10, con los mismos gastos de construcción y de materiales, la segunda vería aumentado su coste en un 40 por ciento más, atendiendo sólo a la inflación y al incremento de los costes financieros. Muchos analistas coinciden en que la reducción del tiempo invertido es cuestión prioritaria.

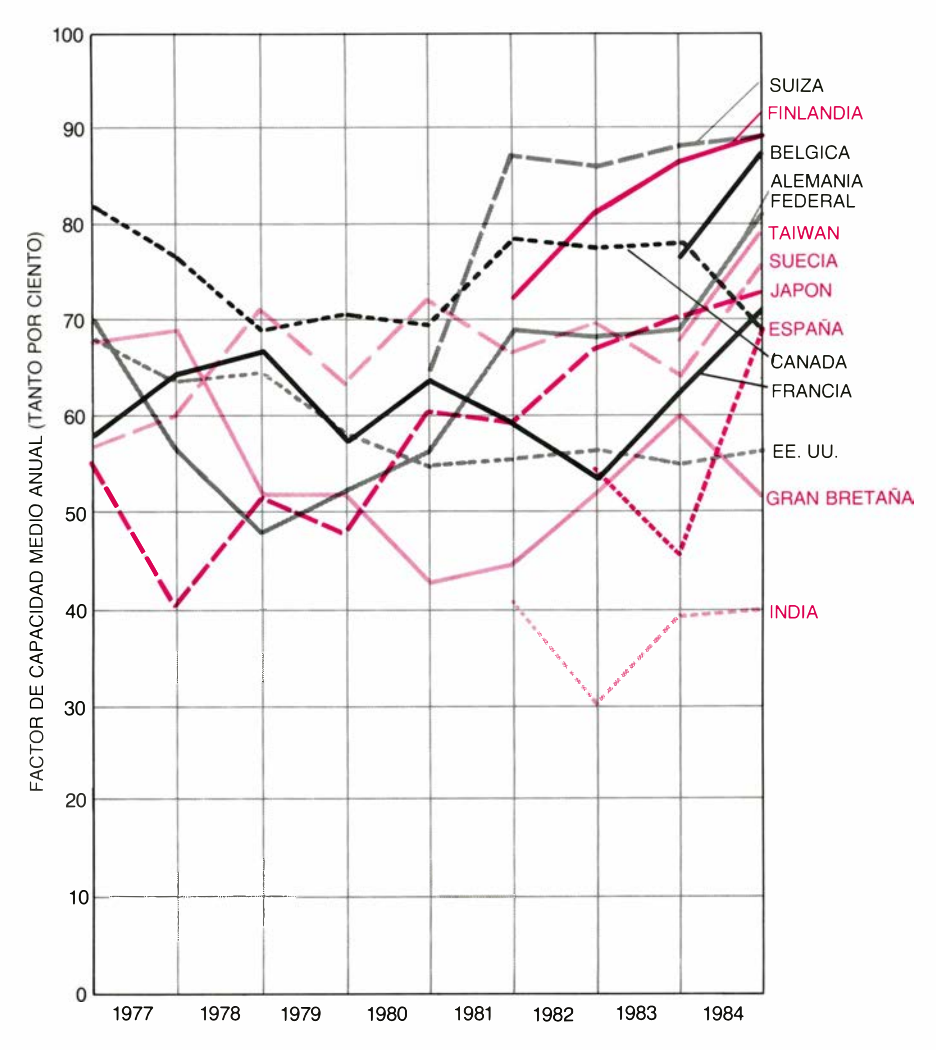
No se ha determinado con suficiente exactitud la parte que cabe asignar al aumento de costes de material, equipo, trabajo y otros afines sobre la inversión total de la central. No obstante, estos costes directos, expresados en dólares por kilowatt, han aumentado en la última década a un ritmo anual del 8 por ciento. En este momento, y no sólo ya por el tiempo de construcción, se está cuestionando que las nuevas centrales nucleares compitan con las de carbón. Las dos únicas centrales cuya construcción durará menos de siete años, y están a punto de terminarse en EEUU, demuestran que una construcción rápida no garantiza unos costes bajos. La unidad de River Bend, en Louisiana, se



1. FORTUNA DISPAR de la energía nuclear en los EEUU. Resulta manifiesta si se compara el número de centrales previstas (*color*) con el número de las que están en funcionamiento (*negro*). Desde los años setenta, momento en que las peticiones de centrales alcanzan su máximo, pocas compañías han solicitado permiso para abrir otras nuevas y muchas han cancelado solicitudes anteriores (la última petición que se ha mantenido en firme se presentó en 1973). No obstante, el número de centrales en servicio ha seguido incrementándose. Las cifras de centrales en operación, de 1986 en adelante (*círculos*), son estimaciones.

convertirá en una de las más caras jamás construidas. Está claro que va a ser necesario algo más que la mera reducción del tiempo para sostener la viabilidad económica de los proyectos nucleares.

¿A qué obedece esa drástica subida de los costes directos? Aunque todo ha contribuido a dicho aumento, el factor trabajo es el que lo ha hecho más espectacularmente. De acuerdo con estimaciones recientes, la mano de obra requerida para la construcción, medida en horas-hombre por kilowatt de potencia, se ha multiplicado, desde comienzos de los setenta, por cuatro; las necesidades de servicios técnicos y de ingeniería se han duplicado. Estos dos tipos de trabajo suponen hoy los dos tercios de los costes directos de la construcción nuclear. El aumento de las necesidades de trabajo puede deberse al bajo rendimiento de este sector laboral en los Estados Unidos. Pero hay más. Se trata, en realidad, de una combinación de factores tecnológicos, de organización, dirección y regulación que es especifica de la industria nuclear.



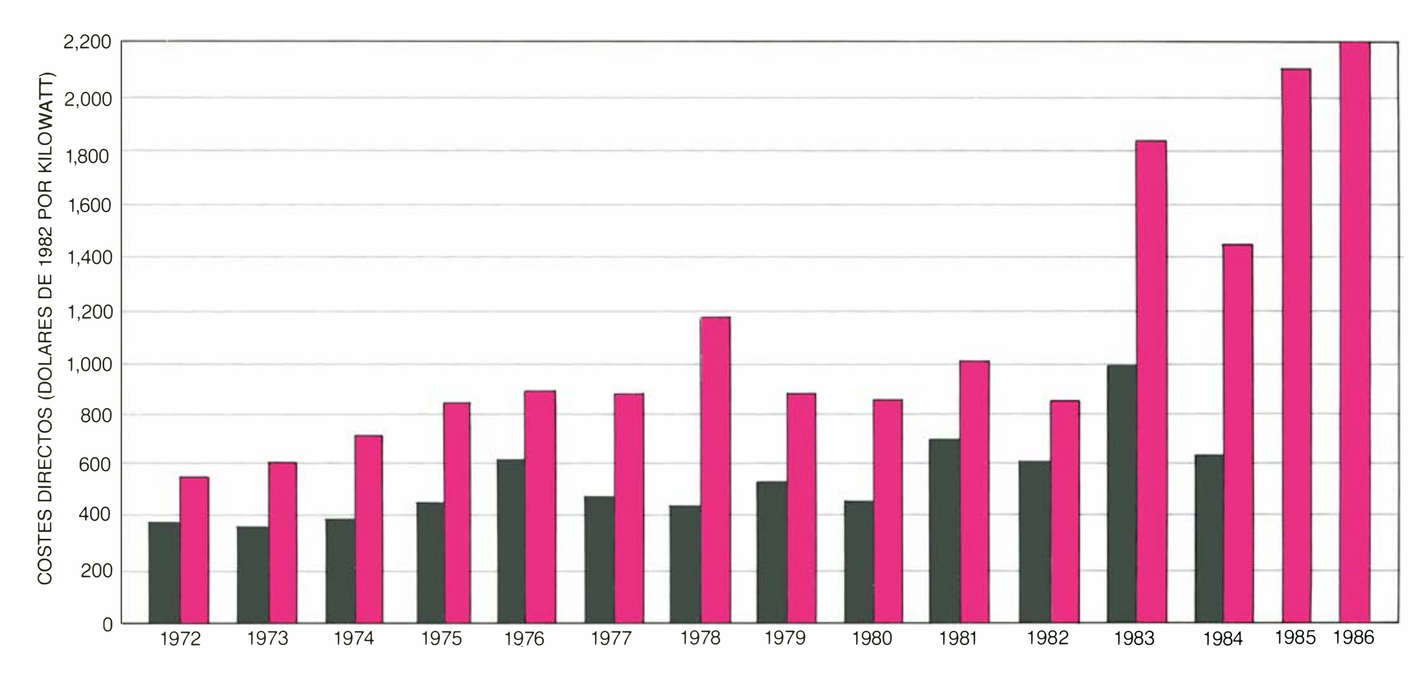
2. FACTORES MEDIOS ANUALES DE CAPACIDAD de las centrales nucleares en distintos países. Indican la energía producida cada año en porcentajes de la potencia nominal. Las causas del relativo bajo rendimiento de la energía nuclear en los EEU no están todavía claras. En parte, es posible que pongan de manifiesto la fragmentación de la industria nuclear, que ha entorpecido la transmisión de información técnica que podría haber mejorado la eficacia, y la ausencia de normalización en la tecnología que limita la aplicabilidad de cualquier mejora técnica. El Instituto para la Operación de Centrales Nucleares (INPO), patrocinado por compañías eléctricas, desempeña un papel importante fomentando una coordinación mayor. Así, algunos reactores estadounidenses han alcanzado, sistemáticamente, factores de capacidad mayores que la media nacional, sacando provecho de una gestión exigente del funcionamiento y mantenimiento.

Las centrales nucleares han ganado en complejidad a lo largo de la última década. El empeño puesto en reforzar la seguridad ha sido el responsable de dicha complejidad; una avalancha legislativa ha invadido, desde principios de los setenta, cuanto se refiere a la construcción y el diseño de las centrales. Se han multiplicado los componentes y los sistemas y se han refinado las revisiones de la construcción y el diseño; los papeles y documentos oficiales requeridos llenan amplios cartapacios. Más aún, la Comisión Reguladora Nuclear (NRC) ha aplicado muchas de las nuevas medidas, de forma retroactiva, a centrales en construcción, obligando a rehacer parte del trabajo.

Ante la poca coordinación que reina en el seno de la industria nuclear norteamericana, no ha sido fácil hacer frente a la creciente complejidad tecnológica. Al igual que en la construcción de centrales de combustible fósil, el diseño y la ingeniería de una central se divide entre diversas organizaciones. El suministrador del reactor es su responsable también, así como de otros sistemas y componentes principales. De un arquitecto-ingeniero depende el grueso restante de la central. Un tercer contratista lleva a cabo la construcción real de la planta. La coordinación de diversas organizaciones independientes (cada una con diferentes términos contractuales y un alcance técnico limitado) en una época de rápido cambio tecnológico y de normativas pone a prueba al jefe de proyecto más experimentado.

Falta de trabazón que caracteriza también a la industria eléctrica: pocas compañías han sabido asimilar la experiencia necesaria. El programa nuclear de casi la mitad de las 55 empresas que poseen unidades nucleares en funcionamiento, o en fase de construcción, consta solamente de una central. Peor todavía, una compañía que construya una segunda central quizá no acierte a sacar partido de su propia experiencia si, como suele suceder, trabaja con un suministrador del reactor o un arquitecto-ingeniero diferente. En el programa nuclear estadounidense han participado cuatro suministradores de reactores y 12 arquitectos-ingenieros; varias compañías eléctricas han contado con su arquitecto-ingeniero propio. Sólo en contados casos el mismo grupo ha colaborado en sucesivas unidades.

No es despreciable la importancia que la experiencia tiene en los costes directos de las actuales centrales nucleares. En razón de ella, se han llegado a producir diferencias de un factor cuatro en dichos costes. Las centrales menos caras las construyó una compañía eléctrica, un suministrador y un arquitecto-ingeniero que colaboraron con anterioridad en varios proyectos. La mayoría de unidades más caras corresponden a empresas inexpertas en proyectos nucleares; las restantes son producto de primeras colaboraciones entre compañías y suministradores.



3. COSTES DIRECTOS DE LA CONSTRUCCION de centrales de carbón (*negro*) y centrales nucleares (*color*) que han ido entrando anualmente en servicio desde 1972 hasta 1984 (en dólares de 1982) por kilowatt de potencia. Se muestran, asimismo, los costes directos previstos para las centrales nucleares que entren en operación en el bienio 1985-86. Los costes directos son una medida de los gastos de materiales y trabajo requeridos en la construcción, pero no incluyen los costes de financiación. Los problemas institucionales y la enorme complejidad que caracteriza a las centrales actuales son las causas de que los costes directos resulten bastante mayores que los invertidos en las centrales de carbón, compensando, significativamente, el menor coste del combustible nuclear.

Los costes directos se han visto incrementados, además, por la falta de estandarización tecnológica y la fragmentación de la estructura industrial. Existen casi tantos diseños cuantas centrales nucleares. La experiencia francesa, donde la compañía nacional y un sólo consorcio suministrador han seguido la política de construir grandes series de centrales muy similares, ha evidenciado la cantidad de beneficios que se han perdido en los Estados Unidos. Los costes de diseño e ingeniería son menores cuando se distribuyen entre muchas centrales. Ingeniería y construcción ganan más experiencia y aumenta, por tanto, su productividad. Los suministradores pueden sacarle un mayor rendimiento a su capacidad de fabricación y entrenamiento. Aunque sean pocas las industrias nacionales de otros países que caigan en el monolitismo francés, en ninguna de ellas se llega al grado de atomización existente en los EEUU, tanto en los proyectos individuales como en las líneas generales del sector.

Esa fragmentación y la falta de normalización tecnológica han complicado también el trabajo de la NRC, que ha tenido que enfrentarse a compañías de diversa capacidad y centrales de características diferentes. La NRC acarrea, además, una serie de problemas que dimanan de su propia estructura. Varios comités investigadores, entre ellos la comisión Kemeny, que analizó el accidente de Three Mile Island, han llegado a la conclusión de que la NRC puede convertirse en un organismo absolutamente ineficaz. Sus dificultades internas han contribuido a la inclusión de trámites carentes de coherencia y arbitrarios en relación con la concesión de licencias, con lo que se ha aumentado el coste y la incertidumbre financiera. Fuentes oficiales de la industria han expresado, en repetidas ocasiones, que hasta que los riesgos financieros asociados a la concesión de nuevas centrales no puedan disminuir, no tiene sentido pensar en nuevos permisos.

Si un equipo competente de ingenieros, con la suficiente información acerca de la solvencia económica y las preferencias de las compañías, de las tendencias reguladoras y de la opinión pública, se propusiese desarrollar una central nuclear a la medida del mercado eléctrico de los EEUU, ¿en qué se parecería a los LWR actuales? Quizás en muy poco. En tres aspectos fundamentales se alejaría su diseño del habitual.

En primer lugar, la potencia del nuevo reactor distaría bastante de alcanzar los 1000 megawatt. Las compañías construyen grandes centrales con la esperanza de ahorrar en idéntica proporción. Los resultados han sido poco convincentes. Según estudios recientes, los inconvenientes derivados de la complejidad de una central de más de 800 megawatt superaría sus posibles ventajas económicas. Las centrales de gran tamaño presentan, además, otro inconveniente: reducir las oportunidades de que una compañía gane experiencia y se beneficie de ella en la construcción de otras centrales menores.

La situación actual de regulación económica desalienta la realización de nuevas inversiones en centrales de gran tamaño. En cualquier caso, la mayoría de las compañías norteamericanas son demasiado pequeñas y su crecimiento demasiado lento para pensar fácilmente en centrales de 1000 megawatt. Tales centrales hacen más difícil que el aumento de la capacidad de una compañía se acomode al crecimiento de la futura demanda, en particular ahora que dicho crecimiento se está volviendo cada vez menos predictible. En teoría, la posesión compartida por varias compañías de grandes centrales nucleares podría remediar dichos problemas; sin embargo, la experiencia enseña que lo que en principio parecería un medio para repartir los riegos, puede convertirse en un mecanismo para aumentarlos, sobre todo si las compañias están, cada una, sujetas a diferentes comisiones reguladoras estatales.

En un futuro, la competencia en el mercado eléctrico entre compañías ya establecidas probablemente aumentará, y también lo hará entre compañías y nuevos productores de energía: generación de electricidad y calor para procesos industriales. Dejando aparte

las compañías de probada competencia y eficacia, que podrían preferir crecer mediante grupos menores de 1000 megawatt, los mayores riesgos comerciales servirán para que disminuyan los encargos de grandes centrales nucleares.

El segundo aspecto diferenciador entre el diseño hipotético y la práctica actual concierne a la construcción de la central. El diseño tendría que permitir que la mayor parte de la fabricación se llevara a cabo en la nave, y no en el punto de emplazamiento. Aumentarían entonces la productividad y calidad del trabajo, puesto que ambos índices suelen ser mayores en el ambiente controlado de una factoría que en planta. La centralización de la producción posibilitaría la aplicación de técnicas avanzadas de fabricación automática y de inspección.

La ventaja quizá más importante que traería dicha centralización sería la normalización del diseño. Esto conllevaría una acumulación de experiencia cada vez mayor, con las consiguientes mejoras en el rendimiento y economía. La limitación en el peso y en el volumen de los componentes que tienen que transportarse haría que la fabricación centralizada privilegiara las centrales de menor tamaño. No obstante, las grandes centrales podrían verse también favorecidas con diseños modulares que permitiesen, además, el ensamblaje en planta.

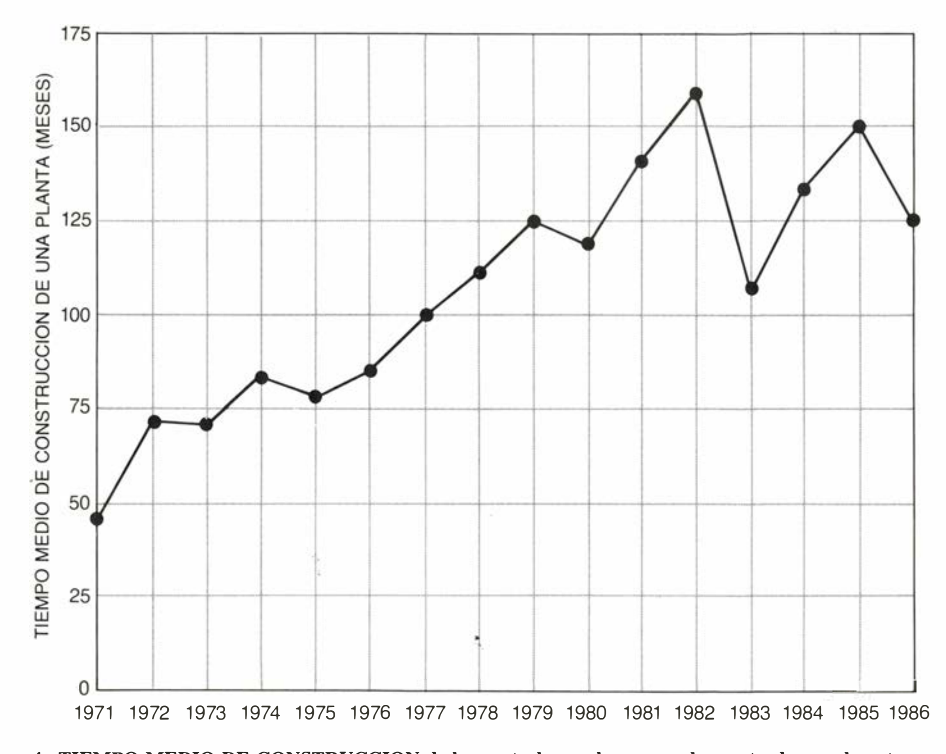
La tercera novedad se refiere a la seguridad. Se trataría de conseguir un reactor que fuese, en comparación con los actuales LWR, más fácil de conceder y emplazar, así como menos sensible a cambios reguladores. El problema implicado en los actuales reactores no es el de su inseguridad; numerosos estudios técnicos han demostrado que dichos reactores suponen una amenaza para la salud pública mucho menor que cualquier riesgo común. El problema está en que la base técnica de dicha conclusión no es clara y su exactitud resulta difícil probar.

La seguridad de un reactor depende de su capacidad para evitar que escapen a la atmósfera cantidades significativas de productos radiactivos. Ya que en el combustible se encuentra la mayor parte de estos productos, sólo el sobrecalentamiento del núcleo debido a un fallo de refrigeración provocaría una liberación importante de radiactividad. Un reactor nuclear está diseñado de tal manera que resulte imposible mantener la reacción en cadena si la temperatura sobrepasa cierto nivel, el reactor se para automáticamente si el núcleo se sobrecalienta. A pesar de ello, los productos de fisión siguen liberando grandes cantidades de calor (calor residual); si éste no se evacuase, la temperatura seguiría aumentando hasta dañar al combustible. Así pues, hay que colocar ciertas barreras (por ejemplo el recinto de contención) que aislen el núcleo del medio ambiente, en previsión de escapes radiactivos procedentes del reactor.

La seguridad de los LWR actuales se basa en una jerarquía compleja de sistemas. Algunos de ellos intentan evitar un sobrecalentamiento accidental del combustible, en tanto que otros previenen la liberación de radiactividad en el caso de un fallo en la refrigeración que dañe el combustible. Algunos dispositivos de seguridad, así el recinto de contención, son pasivos. Pero la mayoría de los sistemas, pensemos en refrigeradores de emergencia que introducen agua dentro del núcleo, tienen que ser activados automáticamente o por personal técnico. Puesto que ni el hombre ni los componentes electromecánicos garantizan una absoluta fiabilidad, se han incorporado múltiples sistemas de reserva para ejercer funciones de seguridad. Dicha incorporación no elimina totalmente la posible liberación de radiactividad, pero disminuye la probabilidad de que ocurra hasta un nivel muy bajo.

La evaluación del riesgo no es asunto sencillo. Hay que identificar todas las posibles secuencias de un fallo, humano o electromecánico, que conduciría a una liberación radiactiva, estimando sus distintas probabilidades y sus consecuencias. Tales análisis probabilísticos del riesgo son ininteligibles para el público inexperto. Además, las liberaciones de radiactividad grandes son tan raras que no es nada fácil verificar los cálculos mediante comparación con experimentos reales (en otros campos como, por ejemplo, la seguridad aeronáutica, donde se aplican los mismos métodos de análisis del riesgo, la situación es diferente). La única base en la que se apoya la gente para juzgar la seguridad de un reactor está en la credibilidad que tenga el investigador o la entidad que haya llevado a cabo las estimaciones. Inevitablemente, cualquier suceso que menoscabe dicha credibilidad (el accidente de Three Mile Island) adquiere una significación que trasciende, con mucho, su impacto técnico.

En parte debido a su novedad y complejidad, las técnicas de análisis probabilístico del riesgo no han desempeñado, en el pasado, el papel predominante que vemos hoy. La normativa se basaba en razones ingenieriles para es-



4. TIEMPO MEDIO DE CONSTRUCCION de las centrales nucleares que han entrado anualmente en servicio, en el período de 1971 a 1986, en los EEUU; ha pasado de menos de cinco años a más de diez, aproximadamente, el doble del necesario para construir una central de carbón o terminar una central nuclear en Francia o Japón. Este dilatado tiempo de construcción de los proyectos nucleares los ha hecho muy sensibles a los efectos de los altos tipos de interés y de la inflación elevada, con lo que el coste total se ha disparado.

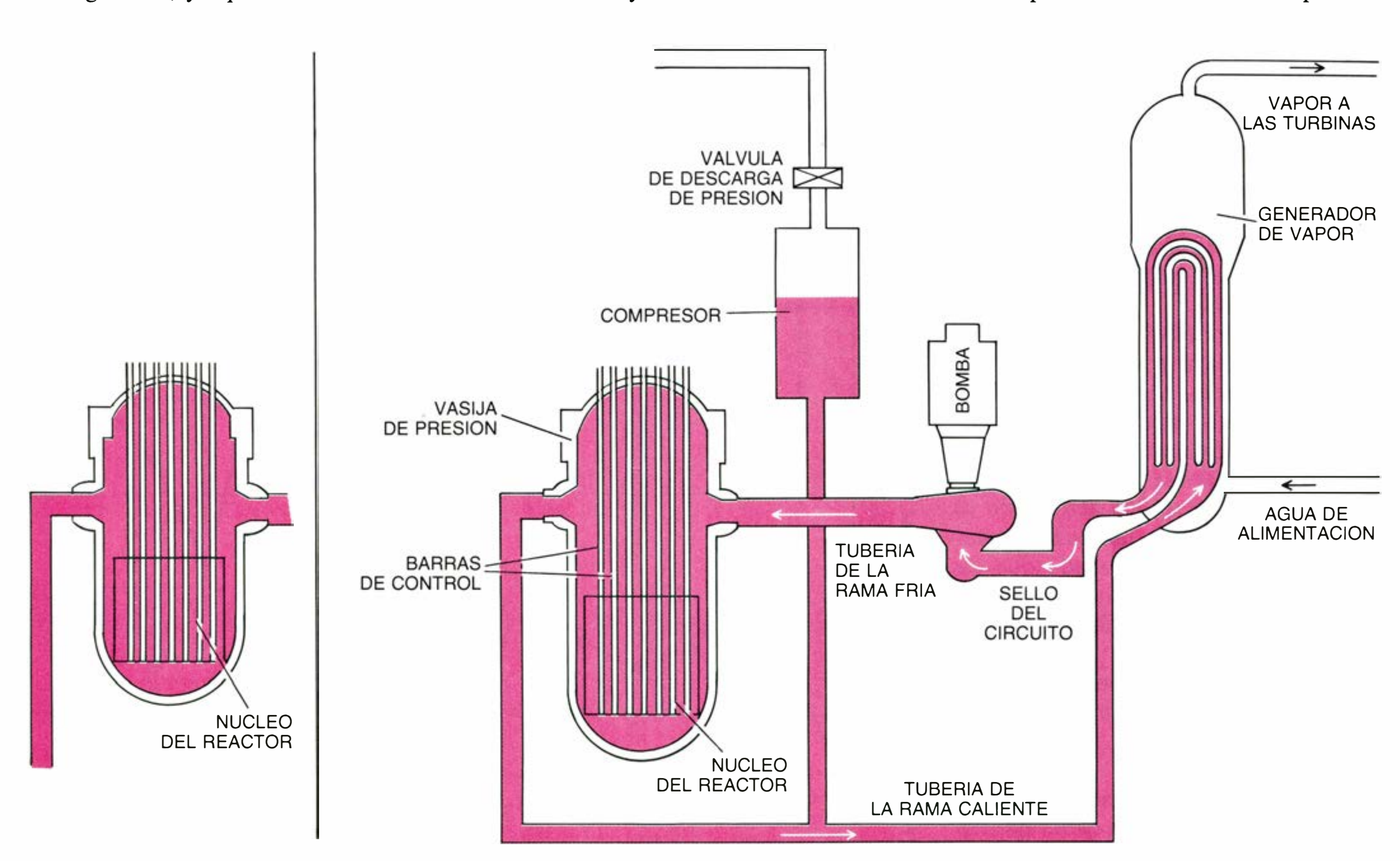
tablecer reglas pormenorizadas de los distintos componentes, y establecía, de acuerdo con todos los requerimientos, una guía de seguridad integral. La ausencia de una base cuantitativa clara que juzgue la validez de las diversas exigencias, ha revestido a la normativa de un aire de imprecisión que ha inquietado tanto a la industria como a sus críticos. Además, la dependencia que la NCR tiene de un número cada vez mayor de normas que, con todo detalle, disponen cómo tienen que diseñarse, construirse y funcionar las centrales, ha enturbiado de tal forma la relación entre la industria y el organismo regulador que los usuarios de la tecnología nuclear han cedido gran parte de la responsabilidad de su gestión a una burocracia reguladora distante.

Para la industria resulta de importancia vital acatar la normativa. En contrapartida, se han agotado su creatividad e iniciativa. No hay apenas lugar ni para corrientes innovadoras en materia de seguridad, ya que se hace más fácil seguir procedimientos aceptados que conseguir la aprobación de otros nuevos. A menos que se aligere el peso de la normativa dominada por la seguridad, será difícil que la energía nuclear recobre la viabilidad comercial.

¿Qué cambios debieran acometerse en el diseño de un reactor para mitigar estos problemas? Una opción prometedora, que ha atraído la atención en los EEUU y en otros países, es la siguiente: un reactor en el que, en caso de un fallo de la refrigeración, el calor generado por los productos de fisión del núcleo se disipe sin problemas mediante procesos naturales de transmisión térmica (por convección o radiación). La refrigeración pasiva bastaría para evitar el sobrecalentamiento y daño del núcleo, con lo que se eliminaría la necesidad de sistemas de refrigeración forzada y trabajadores que los activasen.

Una estrategia como ésta simplificaría el diseño y construcción de una central; en particular, en lo relativo a sus sistemas no nucleares. Dichos sistemas están sujetos hoy a normas reguladoras estrictas, pues sus eventuales fallos podrían provocar un daño serio en el núcleo. Un reactor seguro dejaría que el resto de la central se proyectara y construyera siguiendo las normas de centrales térmicas convencionales, con lo que el ahorro en los costes de equipos y trabajo podría ser notable.

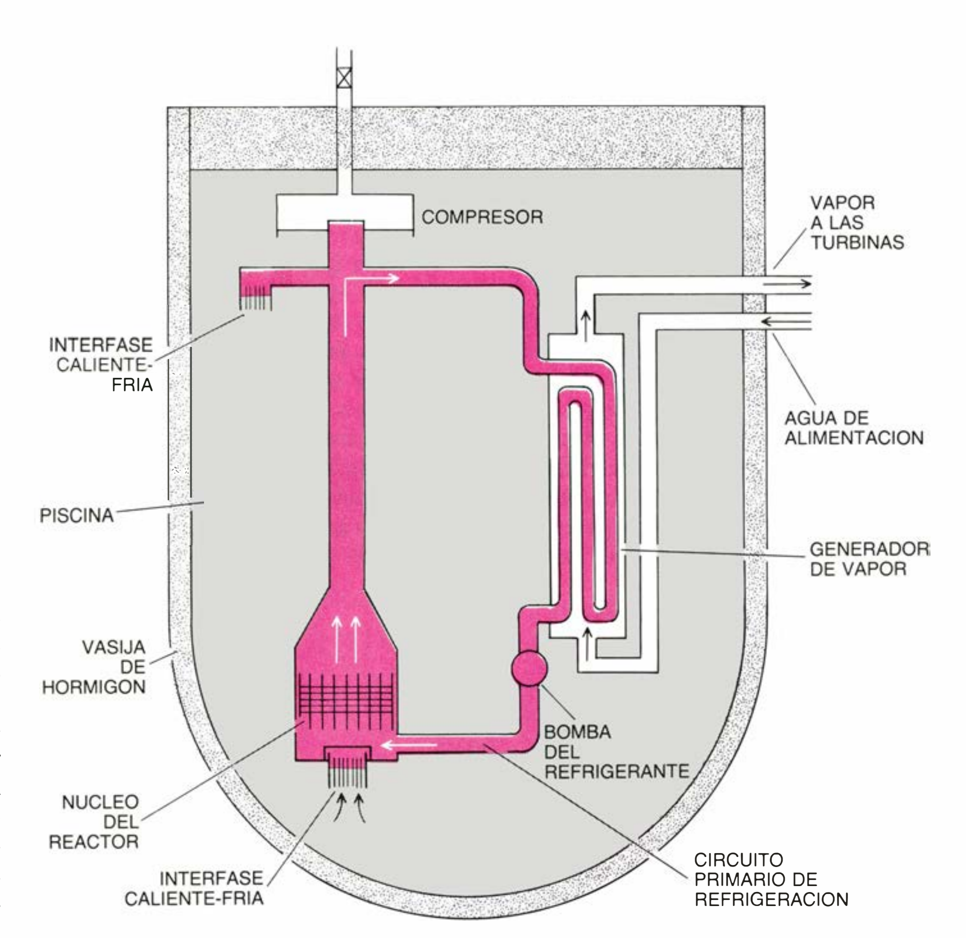
Una seguridad pasiva simplificaría también el análisis de riesgos, tanto por la reducción del número de sistemas a considerar como por la eliminación del problema de predecir la probabilidad y efectos de los errores de operación. Para el público, los conceptos de evacuación pasiva del calor y de seguridad, no dejada tanto en manos del personal de operación, son más asequibles que las complejas características de seguridad de los LWR. Además, la demostración de dichas nociones sería más convincente, con tal de someter a un reactor sin personal a toda clase de pertur-



5. REACTOR AVANZADO DE AGUA A PRESION, en esquema. Comparte su diseño básico con los reactores de agua a presión convencionales. El agua del circuito primario de refrigeración pasa a la vasija de presión a través de la tubería de la rama fría, extrae el calor de núcleo del reactor y arriba al generador de vapor a través de la tubería de la rama caliente; el vapor pasa luego a las turbinas. Un compresor mantiene la presión del circuito primario de refrigeración evitando la ebullición del agua. Un sello de bucle protege a la bomba del refrigerante de burbujas de aire. Este nuevo diseño, producto de la colaboración entre Westinghouse Electric Corporation, Mitsubishi Company, compañías eléctricas y gobierno japoneses, consta de un núcleo mayor, una vasija de presión más profunda que los de los reactores actuales (*izquierda*). El aumento del tamaño del núcleo aligera la densidad de potencia del reactor y, por consiguiente, la densidad de productos de fisión. Por tanto, en caso de un fallo de refrigeración, el calor residual produciría un aumento de la temperatura del núcleo más lento que en los reactores actuales. La vasija de presión, al ser más profunda, sitúa al núcleo en un nivel bastante más bajo que la tubería de la rama fría y el sello del circuito, asegurando, con ello, que la pérdida de refrigerante a través de una pequeña rotura en la rama fría no dejaría el núcleo al descubierto (situación que, en cambio, puede darse en los reactores que están actualmente en servicio). En el caso de cualquier fallo en la refrigeración, el gran volumen de agua existente por encima del núcleo tardaría más tiempo en desaparecer por ebullición que en los diseños actuales, aumentándose, de ese modo, el tiempo del que dispondrían los ingenieros y demás personal técnico para reaccionar ante cualquier emergencia. El reactor, que produciría una potencia algo superior a la de los actuales, incorporaría otras muchas mejoras pensadas para disminuir los costes del combustible, mejorar el rendimiento y facilitar el mantenimiento.

baciones simuladas, ya sean accidentales o intencionadas. Por último, eliminando el riesgo de que el núcleo sufra graves daños, se aliviarían las preocupaciones que futuros propietarios tuvieran acerca de los enormes costes que acarrea hacer frente a un accidente, incluso de consecuencias mínimas para la salud.

Los ingenieros han comenzado a investigar los cambios a introducir en los LWR para disminuir el tamaño, incrementar la fabricación en factorías y aumentar la seguridad pasiva. El Instituto de Investigación de la Energía Eléctrica (EPRI), patrocinado por las compañías eléctricas, ha iniciado un programa para definir los requisitos técnicos de la nueva generación de LWR. El programa del EPRI, que se centra en unidades grandes, subraya la necesidad de diseños simplificados para centrales de reactores fáciles de manejar en caso de emergencia. Si se interrumpe la circulación del refrigerante en un reactor convencional de agua a presión, el agua del núcleo puede hervir muy rápidamente y dejar al descubierto el combustible, con el consiguiente riesgo de sobrecalentamiento. El EPRI está considerando, como uno de los pasos hacia la seguridad pasiva, aumentar la profundidad a que el núcleo está sumergido en la vasija de presión. En caso de emergencia, el agua que cubre el núcleo tardaría más en entrar en ebullición que en los reactores actuales, con lo que el posible daño en el combustible podría retardarse. Dichas medidas no eliminarían la necesidad de unos sistemas de refrigeración de reserva y de otros de contención, pero darían a los técnicos más tiempo que el que ahora tienen para responder ante una emergencia.



6. REACTOR PIUS de máxima seguridad, diseñado por ASEA-ATOM, de Suecia. Consta de un núcleo y de un circuito primario de refrigeración similares a los de un reactor de agua a presión, pero está inmerso en una piscina de agua fría borada. La piscina y el circuito de refrigeración quedan conectados hidráulicamente por interfases estables, pues el refrigerante, que está caliente, se halla situado por encima del agua de la piscina, que es más densa. En su funcionamiento normal, la presión ejercida por la bomba del refrigerante evita que el agua de la piscina entre en el núcleo. En caso de fallo en la refrigeración, la convección natural introduciría el agua de la piscina en el circuito (*flechas negras*). El boro del agua absorbería neutrones interrumpiendo la reacción en cadena, el enorme volumen de agua bastaría para refrigerar el reactor durante varios días, sin que fuese necesaria la intervención de operadores o sistemas de refrigeración de emergencia.

Ahora bien, están surgiendo nuevos diseños de reactores que van más lejos, en el camino de la seguridad pasiva sobre todo. Entre los más prometedores de estos reactores intrínsecamente seguros está un reactor de pequeño tamaño, de alta temperatura y refrigerado por gas (HTGR), desarrollado en Alemania. El reactor tiene un núcleo con la apariencia de un lecho de "cantos rodados" en el que el combustible está embutido en bolas pequeñas de grafito. Dicho combustible soporta temperaturas de hasta 1600 grados centígrados sin liberar productos de fisión. El pequeño tamaño y la gran relación superficie-volumen del reactor asegura que la temperatura del núcleo no supere esos 1600 grados, ni siquiera aunque el reactor perdiese todo el refrigerante de helio mientras operase a plena potencia. La reacción de fisión en cadena se detendría automáticamente tras un aumento inicial de temperatura, de la misma forma que en los LWR. Los productos de fisión continuarían produciendo calor, pero la transmisión pasiva de dicho calor a través de las paredes de la vasija del reactor sería suficiente para estabilizar la temperatura a un nivel seguro. Ya que el pequeño tamaño constituye elemento crucial para conseguir un diseño seguro, dicho reactor no podría producir más de unos 100 megawatt de electricidad, por lo que una central podría constar de varios de estos módulos. En los EEUU se está proyectando un HTGR de tamaño comparable, pero con un diseño del núcleo diferente.

Un modelo sueco (PIUS, del inglés "Proccess Inherent Ultimately Safe") presupone una reconfiguración radical del LWR. El núcleo, el circuito primario de refrigeración y los generadores de vapor están inmersos en una balsa de agua fría y borada encerrada en una estructura de presión de hormigón pretensado. La piscina y el circuito refrigerante primario están conectados hidráulicamente; ahora bien, en condiciones normales, basta la presión suministrada por las bombas de refrigeración para evitar que el agua de la piscina entre en el núcleo. Cualquier perturbación en el sistema de refrigeración descompensaría el equilibrio y el agua borada entraría en el núcleo. El boro, que es un absorbente eficiente de neutrones, interrumpiría la reacción en cadena y el agua fría evacuaría el calor residual. No serían necesarios ni operadores ni ningún dispositivo electromecánico para desencadenar esta serie de acontecimientos.

Un tercer modelo con seguridad pasiva nace del empeño puesto en el desarrollo de reactores reproductores. Estos producen más combustible que el que consumen, ya que transforman isótopos que no sirven de combustible (por ejemplo, uranio) en isótopos fisi-

bles (verbigracia, plutonio). Dichos reactores utilizan sodio líquido como refrigerante, y hasta hace poco se concibieron de tamaños que oscilaban entre los 1200 y 1500 megawatt. El desarrollo de estos reactores de gran tamaño continúa fuera de los EEUU, sobre todo en Francia y URSS. Sin embargo, el esperado aumento rápido de los precios del uranio, que fue el principal motivo económico para el desarrollo de los reactores reproductores, no parece que vaya a producirse. Los ingenieros norteamericanos han comenzado a explorar reactores líquido-metal de un tamaño mucho menor, comparables a los HTGR modulares descritos y capaces de disipar pasivamente, en caso de emergencia, el calor desde las paredes de la vasija. Al igual que los reproductores de gran tamaño, estos reactores producirían cantidades significativas de plutonio, y éste podría recuperarse reprocesando el combustible.

El tener que reelaborar combustible reduce el atractivo de estos pequeños reactores refrigerados por metales líquidos, a la hora de pensar en ellos como una opción distinta de los LWR comerciales. Se trata de una carga a añadir al coste de comercialización de la tecnología del reactor y a la complejidad técnica y organizativa de la producción reactor-combustible. Existe, además, la preocupación generalizada de que la reelaboración del combustible y el reciclado del plutonio favorecería la proliferación de armas nucleares.

7. REACTOR MODULAR DE ALTA TEMPERATURA refrigerado por gas. Utiliza helio a presión (*color*) para transportar el calor del núcleo del reactor al generador de vapor. En este diseño, propuesto por la empresa KWU/INTERATOM, de Alemania, el combustible está constituido por partículas pequeñas de uranio, revestidas individualmente con capas de grafito y carburo de silicio y embutidas en bolas, pequeñas esferas de grafito. Cientos de miles de bolas conforman el núcleo; pueden añadirse continuamente nuevas bolas, mientras se descargan las ya gastadas por la parte inferior de la vasija del reactor, lo que permite que éste pueda recargarse durante su funcionamiento. El combustible puede soportar altas temperaturas sin sufrir daño alguno. Debido a que el reactor es pequeño y, por tanto, tiene una relación superficie-volumen alta, en caso de fallo del sistema de refrigeración la temperatura del núcleo se mantendría en un nivel estable; en efecto, la refrigeración por aire y radiación térmica extraerían el calor al exterior de la vasija. Un reactor así produciría unos 100 megawatt de electricidad; una central podría combinar varios módulos como éste.

Muchos industriales del sector nuclear de los EEUU no consideran realista la idea de alejarse demasiado de la tecnología actual. Sostienen que es más acertado mejorar progresivamente la técnica, aprovechando la experiencia adquirida con los LWR, que cambiar a una concepción nueva. Cualesquiera que sean las ventajas de los nuevos sistemas, se dice, van a encontrarse en la práctica con una gran cantidad de problemas imprevistos, como les sucedió a los LWR. Las críticas cuestionan también que fuera acertado para los EEUU apartarse de la tendencia internacional. Francia, Alemania, Japón y otros países están comprometidos con la tecnología de los LWR y han dedicado muchos recursos a su mejora. Participando en proyectos internacionales, como lo están haciendo dos suministradores de reactores estadounidenses, con socios japoneses, la industria norteamericana podría beneficiarse de los avances extranjeros, aunque siguiera encogida la demanda interior.

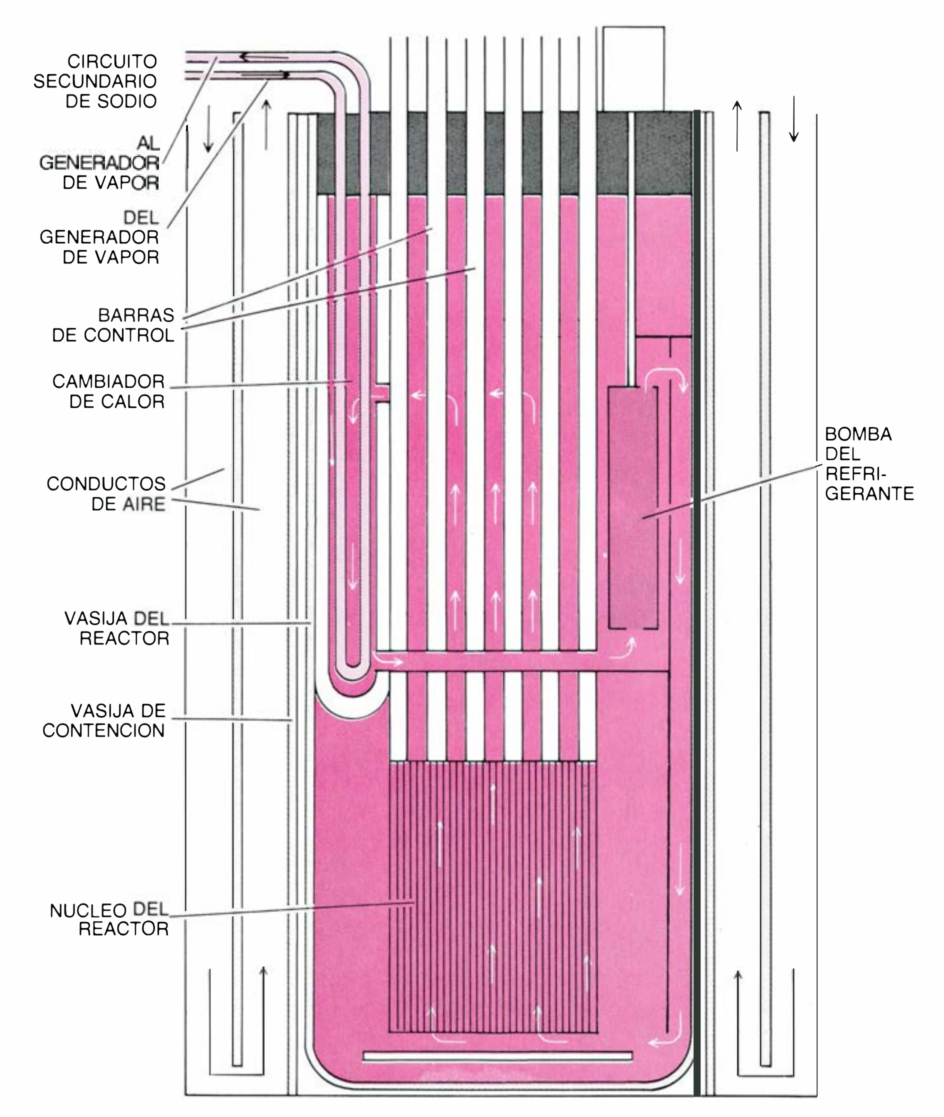
Más razonable que adoptar una tecnología radicalmente nueva sería, aducen, concentrarse en desmantelar las barreras institucionales que se oponen a que los LWR alcancen la misma viabilidad económica en los EEUU que en otras partes. Entre esos objetivos: robustecer la previsión de las normas reguladoras de seguridad y reformar una industria desorganizada y fragmentada. Una mayor normalización, en el proyecto del reactor y de la central, así como en los procedimientos de construcción, tendría que acompañar a las fases organizativa y de gestión. Algunas de las propuestas específicas abogan por una amplia reforma institucional: la creación de compañías regionales de operación de centrales nucleares e incluso de un ente nuclear federal. Otras contemplan la consolidación de la industria nuclear y la asunción por parte de los suministradores de una mayor proporción del riesgo financiero que supone la construcción de una central.

El debate puede resumirse en una cuestión estratégica fundamental: ¿qué es más efectivo, intentar reestructurar las instituciones políticas, industriales y reguladoras para que se adapten a las especiales demandas de la tecnología actual, o acomodar la tecnología a la capacidad, limitaciones y necesidades

de las instituciones existentes en la actualidad?

Si atendemos al primer planteamiento, el problema estriba en que las propuestas más moderadas parecen insuficientes, mientras que los esquemas más radicales implicarían una organización nuclear centralizada, más cercana a los programas extranjeros que al resto de la industria eléctrica estadounidense. La industria nuclear no puede llegar tan lejos en la emulación de sus homólogos extranjeros de mayor éxito. El escaso peso que la planificación y la toma de decisiones centralizadas tienen en la economía de los EEUU complica el desarrollo de la energía nuclear, pero es reflejo de unas preferencias sociales que probablemente se desea que persistan.

La relativa abundancia de otras fuentes propias de energía distingue a los EEUU de otros países que están cosechando éxitos con la energía nuclear. Así pues, la energía nuclear continuará siendo allí, no un imperativo económico y político, como en otras partes, sino una más de las fuentes de energía existentes. Cualquier intento de conceder a la energía nuclear un trato preferencial o de sacarla del terreno de la competitividad no tendría, posiblemente, mucho respaldo. El programa nuclear de los EEUU debe partir de la base de que el país cuenta con una de las sociedades más abierta y beligerante; cualquier maniobra por proteger de la opinión pública a sus instituciones industriales y reguladoras no tendría posibilidades de éxito. Al final, la forma más segura de reanimar la energía nuclear en los EEUU puede ser el desarrollo de una nueva generación de productos mejor adaptados a las condiciones del mercado.



8. REACTOR MODULAR refrigerado por metal líquido: tiene múltiples circuitos de refrigeración. Una bomba situada en el interior de la vasija fuerza la circulación del sodio líquido del circuito primario de refrigeración (*color oscuro*) a través del núcleo. En los intercambiadores, el sodio cede calor a los circuitos secundarios de sodio (*color claro*), los cuales se dirigen a los generadores de vapor. Si fallaran las bombas del refrigerante, el sodio continuaría circulando a través del núcleo del reactor por convección natural. El pequeño tamaño del reactor permitiría su enfriamiento mediante la convección de aire a través de las conducciones que rodean la vasija de contención, evitando dañar el núcleo con la consiguiente liberación de radiactividad. Este diseño produciría alrededor de 140 megawatt de electricidad. Al igual que en los grandes reactores reproductores refrigerados por metal líquido, generarían cantidades sustanciales de combustible nuclear mediante la transformación, por captura neutrónica, de uranio 238 en plutonio.

El replanteamiento de las opciones tecnológicas se enfrentaría, por sí misma, a problemas prácticos. Nuevas opciones distintas de los LWR, incluso las que están en una fase adelantada de desarrollo, no llegarían a ser comerciales antes del próximo siglo. Los altos costes y riesgos de comercialización de una nueva tecnología disuadirían a la industria privada sin un claro apoyo gubernamental. Ahora bien, con la mayor parte de las industrias más importantes comprometidas con la tecnología convencional de los LWR, el gobierno podría encontrarse, no ya apoyando, sino dirigiendo el esfuerzo para el desarrollo de los nuevos reactores. El déficit presupuestario y la falta de entusiasmo público por la energía nuclear descartan, por el momento, esa posibilidad.

En cualquier caso, el intento sería prematuro. La experiencia reciente advierte que, si el gobierno pretendiere optar por el desarrollo de una determinada tecnología, podría correr el enorme riesgo de que la industria no quisiera utilizarla. Un requisito indispensable para llevar a cabo un esfuerzo serio en el desarrollo de una nueva generación de reactores es que exista, por parte de las compañías eléctricas, un compromiso firme de investigación. Dicha esfuerzo debería esperar, probablemente, reformas reguladoras, cambios institucionales y mejoras en los LWR, y aguardar a que se aclarase el porvenir de la tecnología actual.

Entre tanto, gobierno e industria harían bien en marcarse un objetivo más modesto: elegir, en los próximos tres o cuatro años, una de las prometedoras tecnologías de seguridad pasiva para desarrollar, como prueba conceptual, un dispositivo experimental que sirviese, antes que como demostración comercial, de ensayo de su viabilidad tecnológica. Dicha planta podría demostrar el principio de la seguridad pasiva a un coste bastante moderado. Podría servir de medio para resolver cuestiones cruciales sobre regulación normativa que afectara a los sistemas pa-

sivos de seguridad. La experiencia obtenida en el proyecto ayudaría a tomar la decisión sobre la conveniencia de una planta de demostración comercial. Si ahora se llevase a cabo dicho proyecto, podría tomarse una decisión de ese tenor a mediados de los años noventa, cuando el futuro de los LWR convencionales esté más claro que en la actualidad.

El coste de dicho proyecto podría correr a cargo del gobierno y un consorcio de compañías eléctricas y suministradores, incluidos proveedores extranjeros que estuviesen interesados. El interés de los suministradores y compañías eléctricas a la hora de compartir los costes de desarrollo serviría para determinar cuál de los diseños conceptuales debería seleccionarse. El proyecto podría llevarse a cabo con un presupuesto anual inferior al que el gobierno dedica al desarrollo de los reactores reproductores.

Así pues, el futuro previsible de la energía nuclear en los EEUU se despliega en tres posibilidades. Primera: la vuelta a la tecnología de los LWR en una forma mejorada, quizás antes de diez años. El resurgimiento de la tecnología tendría lugar en el seno de una industria más eficaz, con organizaciones menos numerosas y más competentes operando en un clima normativo más estable. Los suministradores extranjeros de LWR que, entre tanto, contarían con la experiencia de mercados nacionales más dinámicos, podrían introducir muchas de las mejoras y figurar comercialmente en dicho resurgimiento. En un segundo “escenario”, los LWR tradicionales no recuperarían la aceptación comercial; ahora bien, pasados unos años, una segunda generación de reactores de pequeño tamaño, fabricación modular centralizada e intrínsecamente seguros permitiría que la energía nuclear se convirtiese en la mayor fuente de electricidad en el próximo siglo.

Ante ese par de “escenarios” no podemos apostar todavía por cuál sea el más probable; tampoco hay necesidad de hacerlo. El empeño puesto en la reforma de las instituciones y en la mejora de la técnica habitual deben seguir a la par que las investigaciones en técnicas totalmente nuevas. De no avanzar en alguna de esas dos direcciones, se desembocaría en una tercera: la desaparición de la energía nuclear como una opción para el futuro. Esto representaría una pérdida que la nación difícilmente podría permitirse.

Giroscopios ópticos

Aunque carecen de partes móviles, detectan movimientos de rotación lo mismo que los giroscopios mecánicos. Tales instrumentos han adquirido gran importancia en los sistemas de guía para la navegación

Dana Z. Anderson

La palabra giroscopio nos trae a la memoria el recuerdo del trompo o peonza que, impulsado por un cordel, es capaz de mantenerse en equilibrio incluso sobre la punta de un lápiz. Tal juguete constituye una primaria representación de una familia de instrumentos cuya estructura esencial la constituye un disco en rápida rotación. El momento angular del disco impide que éste cambie de orientación, aun cuando gira todo el aparato. En virtud de ello, el giroscopio puede mostrar la magnitud de una rotación y, de este modo, facilitar información direccional para la navegación. De hecho, casi todos los vehículos algo más complicados que el automóvil, e incluso algunos que lo son menos, dependen de los giroscopios para mantener su trayectoria. En este orden, los giroscopios constituyen el núcleo de los sistemas de guía por inercia de aviones y buques transoceánicos. Ello se debe a que tales instrumentos funcionan sin necesidad de estímulos exteriores. El giroscopio detecta el movimiento incluso dentro de un recinto cerrado que impida tanto la observación de las estrellas como la acción del campo magnético sobre la brújula.

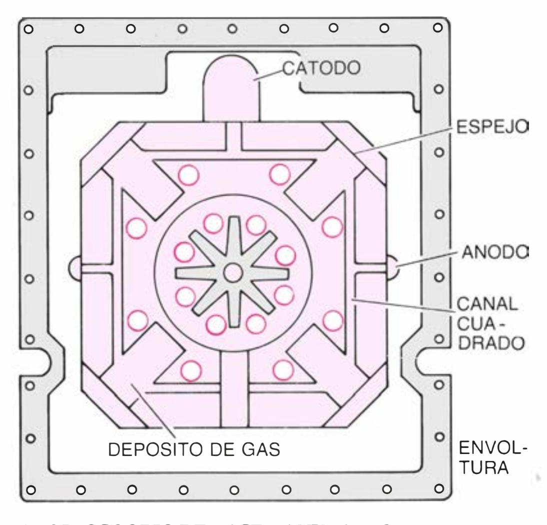
A principios de siglo, los investigadores comprobaron que la luz podía presentar un comportamiento giroscópico: el tiempo que emplea ésta en recorrer una trayectoria circular depende de que el circuito permanezca estacionario o esté girando. La diferencia de tiempo entre uno y otro caso puede utilizarse para medir la velocidad de rotación. Sin embargo, el desarrollo en la práctica de un giroscopio óptico tuvo que esperar al advenimiento del láser y otros adelantos en tecnología óptica, tales como los cables de fibra óptica y espejos de gran poder reflectante. Basados en esos avances se han desarrollado dos clases de sensores ópticos de rotación: los giroscopios de fibra y los giroscopios de láser anular.

El más perfeccionado de estos dos dispositivos es el giroscopio de láser anular. Constituye un instrumento fascinante que permite ya aplicaciones prácticas en sistemas de guía. En los nuevos aviones comerciales, citemos los Boeing de las series 757/767 y cierto número de Airbus A310, se emplean los giroscopios de láser anular en vez de los de tipo mecánico. Aunque la sensibilidad de los nuevos instrumentos es extraordinaria, el uso creciente de giroscopios ópticos en la navegación no responde a una necesidad de mejorar la precisión. Los giroscopios mecánicos también son precisos, pero el que posean partes móviles motiva que sean más complejos que los dispositivos ópticos, que teóricamente carecen de ellas. A pesar de que en la práctica los dispositivo ópticos contienen alguna parte móvil, son más fáciles de mantener, y resultan por ello más baratos, que sus contrincantes mecánicos.

Para comprender cómo aprovechar la luz para medir una rotación, consideremos un observador fijo en un anillo circular. Supongamos que el observador emite un destello luminoso, de manera que la mitad recorra el anillo en un sentido y la otra mitad en el opuesto. Si el anillo no gira, es evidente que ambas mitades deben volver al observador simultáneamente. Pero si gira, por ejemplo en sentido contrario al del movimiento de las agujas del reloj, el punto origen del pulso se moverá hacia el destello que viaja en sentido horario y se alejará del que viaja en sentido antihorario. Por tanto el observador encontrará antes el destello que se mueve en sentido horario que el que lo hace en sentido opuesto. La diferencia entre los tiempos de llegada es directamente proporcional a la velocidad de rotación del anillo, y no influye en ello el que éste gire o no alrededor de su centro.

La diferencia entre los tiempos de circulación se atribuye a una diferencia de longitud entre los caminos recorridos. En el ejemplo anterior, el camino del destello que va en sentido antihorario es más largo que el seguido por el destello que se mueve como el reloj. La diferencia de longitud en los recorridos producida por la rotación constituye el llamado efecto Sagnac, en honor de Georges Marc Marie Sagnac, quien en 1913 descubrió este fenómeno.

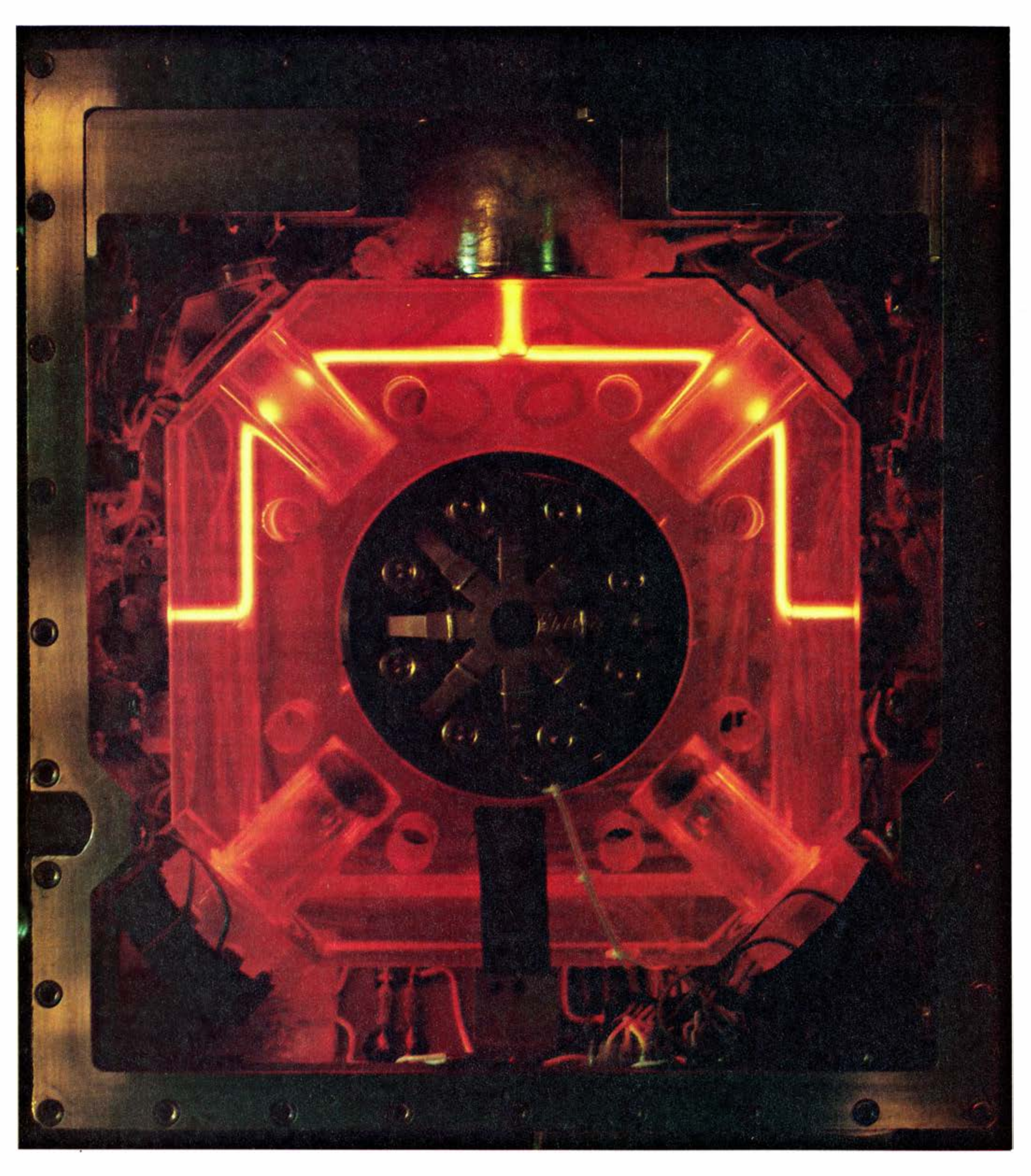
El efecto Sagnac aparece con mayor claridad en el giroscopio de fibra. Por ello estudiaremos en primer lugar este giroscopio, aunque su desarrollo empezó con un retraso de 12 años respecto al giroscopio de láser anular, y todavía no se ha conseguido fabricar con esta



1. GIROSCOPIO DE LASER ANULAR. Como se aprecia en el ejemplar de la página siguiente, consta de un solo bloque de vidrio. El canal cuadrado perforado en el bloque contiene una mezcla de gases. Una corriente eléctrica entre el cátodo, o electrodo negativo, y el ánodo, o electrodo positivo (*véase el esquema de arriba*), excita los átomos del gas y motiva que emitan luz. Los espejos colocados en las cuatro esquinas del bloque mantienen la luz dentro del canal cuadrado, con lo que se forma una onda estacionaria. (La luminosidad de la mitad superior del canal cuadrado es un fenómeno llamado descarga en el plasma; la onda estacionaria no puede verse.) La cinta negra de la parte inferior impide ver detalles que los fabricantes, Litton Guidance and Control Systems, consideran que son de su propiedad.

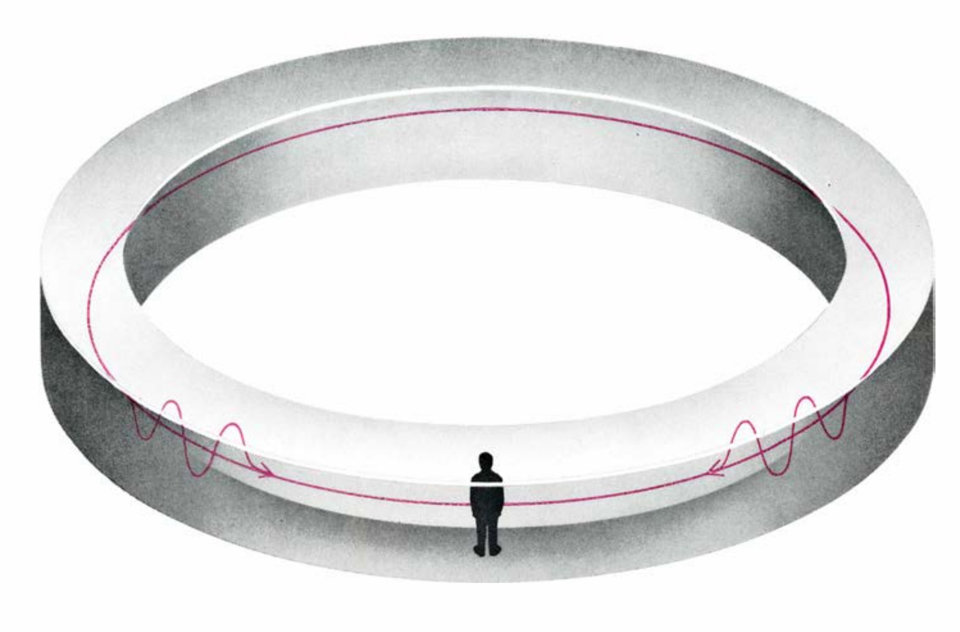
técnica un instrumento práctico verdaderamente sensible a la rotación. El éxito, sin embargo, parece estar a la vista. La parte esencial del giroscopio de fibra es una larga fibra óptica, de medio a un kilómetro de longitud; para reducir el tamaño está bobinada. La fibra actúa como una tubería de luz, quedando ésta confinada dentro de sus paredes. La razón de emplear una fibra tan larga se debe a que la diferencia entre las longitudes recorridas, y por tanto entre los tiempos, aumenta con la longitud de la fibra y con ello se facilitan las medidas de rotación.

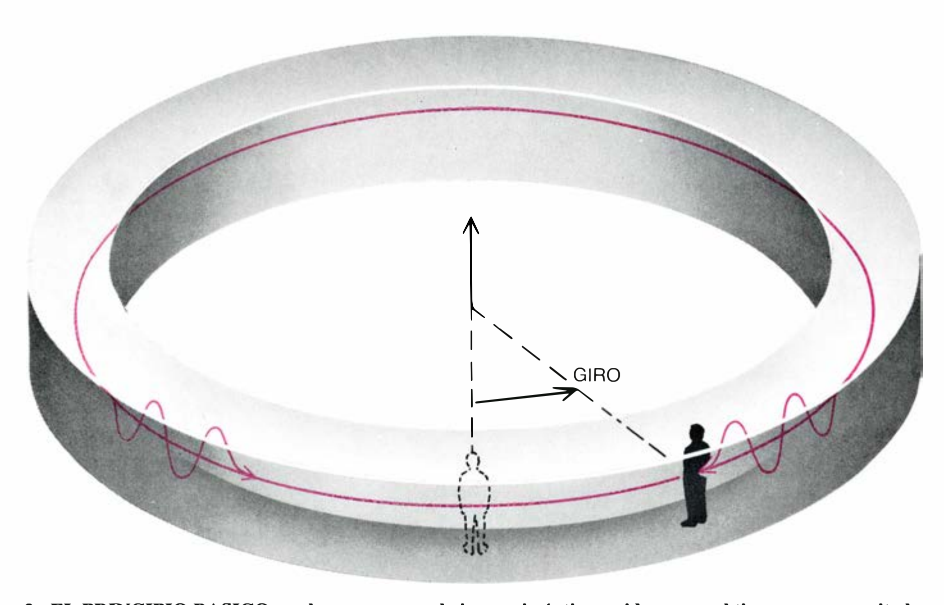
En el giroscopio de fibra se emplea el valor extremadamente bajo de la longitud de onda de la luz para medir con precisión la diferencia entre las distancias recorridas por dos rayos que se propagan en direcciones opuestas. Estos rayos se obtienen mediante un espejo semitransparente que desdobla en partes iguales un haz procedente de la fuente de luz (se emplea casi siempre un láser, aunque no necesariamente). Los dos haces producidos circulan en direcciones opuestas por el carrete de fibra bobinada y al salir se recombinan en el espejo semitransparente. Si se coloca una pantalla a la salida del espejo no se observa ninguna señal luminosa cuando el carrete de fibra permanece



estacionario. Ello se debe a que los dos rayos siguen trayectos de igual longitud y el espejo semitransparente introduce un cambio de fase: el rayo reflejado por el espejo está exactamente en oposición de fase con el transmitido por el espejo. A la salida, por tanto, se produce una interferencia destructiva entre los dos rayos: éstos se anulan mutuamente.

El comportamiento es distinto cuando el carrete de fibra gira. (Para ello se supone que el carrete, la fuente de luz, el espejo semitransparente y la pantalla giran de modo solidario; el carrete no gira respecto a los otros componentes.) Por ejemplo, supongamos que esos componentes estén montados en un avión y que el aparato cambia de rumbo bruscamente. En este caso, como los dos haces se propagan en sentidos opuestos en la fibra recorren longitudes ligeramente distintas. A la salida del espejo semitransparente ya no se cancelan mutuamente, y aparece una mancha brillante en la pantalla. Si el conjunto formado por el carrete y los otros componentes gira con suficiente rapidez, de modo que la diferencia de longitudes entre los dos trayectos recorridos equivalga a media longitud de onda de la luz, la mancha luminosa adquirirá la intensidad de la fuente de luz original. Hemos descrito una versión esquemática del giroscopio de fibra. En la práctica se emplean refinados dispositivos electrónicos para analizar la salida del espejo semitransparente y obtener con ello la velocidad de rotación.





2. EL PRINCIPIO BASICO en el que se apoya el giroscopio óptico reside en que el tiempo que necesita la luz para recorrer una trayectoria circular depende de que la trayectoria sea estacionaria o esté girando. Supongamos que un hipotético observador situado en un anillo estacionario emite un destello luminoso y que dos mitades del mismo se propagan a lo largo del anillo en sentidos opuestos (*arriba*). El observador recibe simultáneamente los dos pulsos. Sin embargo, si el anillo gira, el observador se mueve hacia uno de los destellos y en sentido opuesto al otro recibe en instantes distintos los dos pulsos (*parte inferior*). La diferencia de tiempos es proporcional a la velocidad de giro del anillo. (Los esquemas son de Ian Worpole.)

En 1958 se publicó un trabajo de Arthur L. Schawlow y Charles H. Townes en el que se establecieron los principios básicos del láser. Siguieron poco después los primeros trabajos sobre el giroscopio de láser anular. Clifford V. Heer, de la universidad estatal de Ohio, pronto se dio cuenta de que una cavidad resonante podía ser empleada para medir velocidades de rotación. Una cavidad resonante es un recinto hueco que sirve para reforzar ondas sonoras o bien ondas electromagnéticas. En realidad un láser constituye una cavidad resonante. Consta, a grandes rasgos, de un tubo largo, recto, lleno de un medio amplificador, que puede ser un sólido, un líquido o un gas. Cada extremo se cierra con un espejo semitransparente finamente pulido. Conforme la luz va y vuelve de uno a otro espejo su intensidad queda amplificada. La salida del láser acontece cuando parte de la luz atraviesa uno de los espejos.

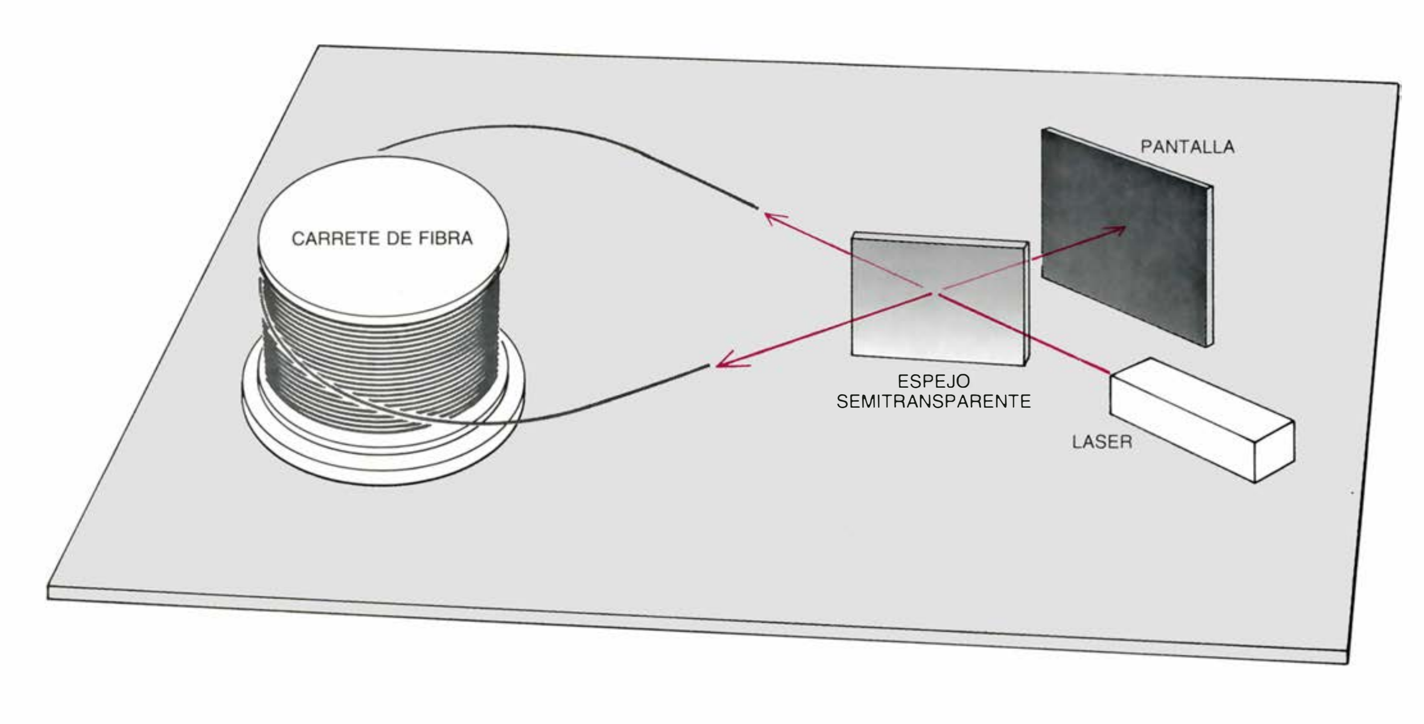
Heer comprobó que formando una cavidad resonante en forma de anillo puede construirse un giroscopio óptico. En el mismo la luz circula muchas veces a lo largo de una trayectoria en lugar de ir y venir entre dos espejos. Heer, e independientemente Adolph H. Rosenthal, en la Kollsman Instrument Corporation, consiguieron ulteriores progresos teóricos. En 1963, Warren M. Macek y Daniel T. M. Davis, Jr., de la Sperry Gyroscope Company, presentaron el primer giroscopio de láser anular. El instrumento estaba formado por cuatro tubos de vidrio, cada uno de los cuales medía un metro de longitud, dispuestos en cuadrilátero. Se hacía circular la luz a lo largo del dispositivo colocando un espejo en cada esquina.

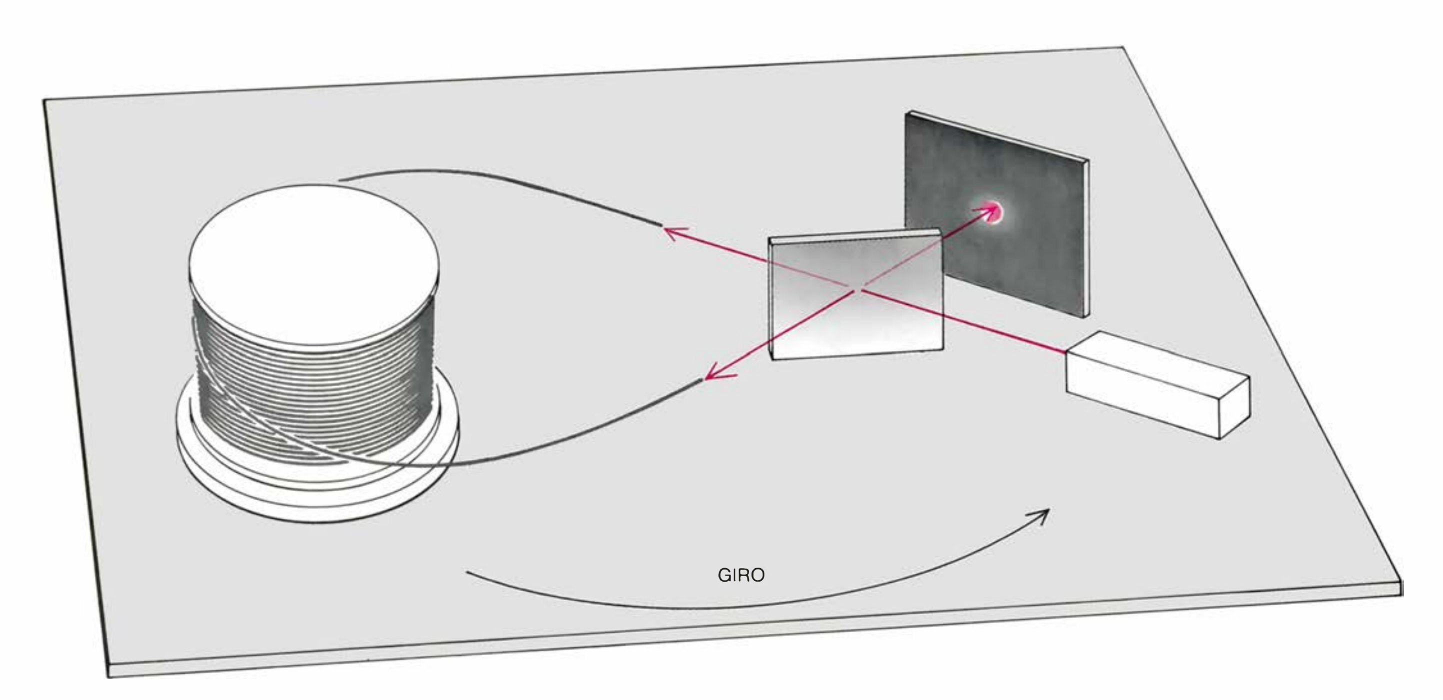
El progreso tecnológico en los años posteriores ha sido espectacular: un giroscopio sensible de láser anular cabe ahora en la palma de la mano. Un giroscopio típico está formado por un solo bloque de vidrio en cuyo interior se ha perforado un canal cuadrado. Este contiene una mezcla de gases, tales como helio y neón. Para completar el láser hay un pequeño número de electrodos y cuatro espejos colocados en el bloque. El cuatro no es un número mágico, pues algunos giroscopios de láser anular presentan un canal

triangular y montan tres espejos. Existen otros con un canal hexagonal y seis espejos.

Aunque en la práctica un giroscopio de láser anular tiene forma poligonal, su funcionamiento se comprende mejor imaginando el caso ideal de un anillo circular. Cuando el anillo permanece estacionario, una descarga en el gas del láser genera, en el interior del anillo, una onda luminosa estacionaria. Al igual que en la cuerda vibrante de un violín, a lo largo de la onda hay una sucesión de nodos (puntos estacionarios) y de antinodos (puntos de máxima oscilación). La longitud de onda de la luz es extremadamente pequeña: unos 0,6 micrometros (un micrometro es una millonésima de metro); por tanto, a lo largo del anillo cabe un gran número de nodos y antinodos. Por ejemplo, una onda en un anillo que tiene un perímetro de 30 centímetros presenta del orden de un millón de nodos y de antinodos.

Supongamos que el anillo gira, debido por ejemplo a que el supuesto avión que lleva el giroscopio cambia de rumbo. La onda estacionaria permanece fija en un sistema de referencia inercial, que no gira. Como consecuencia se manifiesta el efecto Sagnac en el giroscopio de láser anular. Un observador que girase con el anillo vería pasar los nodos y antinodos de la onda es-





3. CON EL GIROSCOPIO DE FIBRA podemos medir una velocidad de rotación detectando la interferencia entre dos rayos de luz que se propagan en sentidos opuestos en una larga fibra óptica bobinada. Los dos haces se obtienen desdoblando un rayo láser mediante un espejo semitransparente. Tras recorrer la fibra, los dos rayos se recombinan en el espejo. Cuando el aparato no gira, los haces se interfieren destructivamente, y no aparece luz a la salida (*pantalla oscura de arriba*). Cuando todo el aparato (láser, el espejo semitransparente, el carrete de fibra y la pantalla) gira, los dos haces ya no se anulan exactamente y se forma una mancha luminosa sobre la pantalla (*abajo*); el brillo de la mancha aumenta con la velocidad de rotación. La velocidad de rotación del aparato, y por tanto del vehículo en el que está montado, se deduce analizando la intensidad de la mancha. Los giroscopios de fibra se encuentran en fase de desarrollo.

tacionaria. El número de nodos que pasarían por el observador sería directamente proporcional al ángulo que hubiese girado el anillo. Contando el paso de los nodos podría medirse el ángulo que hubiera girado el anillo, solidario con el avión.

Un anillo no circular (por ejemplo, triangular o cuadrado) puede emplearse también para detectar la rotación. Sin embargo, en este caso la onda estacionaria no permanece fija en un sistema de referencia inercial cuando el anillo gira. Se observa una rotación a una velocidad más pequeña que la que presenta el anillo. La velocidad de giro de la onda estacionaria depende de la forma del anillo. En lo que concierne al observador, el ángulo de rotación es todavía proporcional al número de nodos que pasan por un punto dado del anillo.

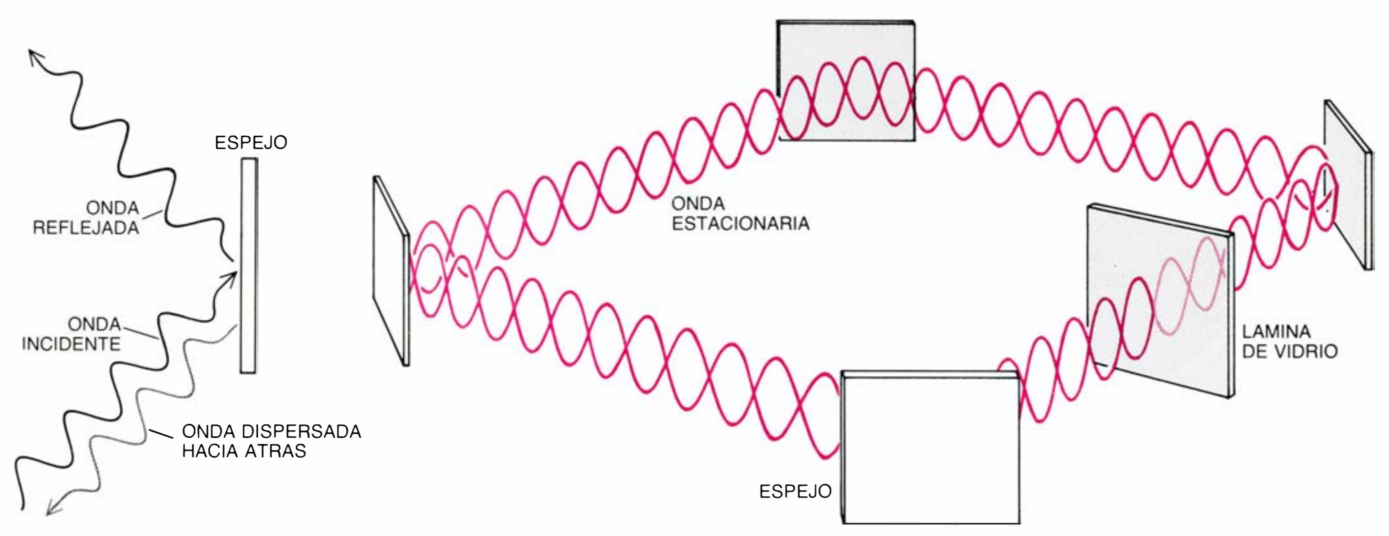
Si el giroscopio de láser anular fuera un dispositivo perfecto, la velocidad con la que pasarían los nodos por un punto dado del anillo sería directamente proporcional a la velocidad de rotación de éste. Si el anillo no girase, la posición de los nodos permanecería fija. En la práctica, sin embargo, intervienen principalmente dos causas de error que motivan una desviación de este comportamiento ideal. Por fortuna, ninguna de las dos es decisiva, pues ambas pueden remediarse.

La primera fuente de error consiste en un efecto de deriva que hace que la onda estacionaria gire incluso cuando el anillo no lo hace. La deriva se origina cuando el gas del interior del anillo fluye. Este flujo aparece como consecuencia del suministro de energía necesario para producir la onda estacionaria. Para aportar la energía se aplica una gran diferencia de potencial entre un electrodo positivo, o ánodo, en un extremo del anillo y un electrodo negativo, o cátodo, en el otro extremo. El elevado potencial ioniza parte del gas y crea un plasma: una especie de "sopa" de electrones y de iones cargados positivamente. Los electrones son atraídos hacia el electrodo positivo y los iones positivos hacia el electrodo negativo. Este flujo induce un movimiento más complicado en los átomos neutros del gas que redunda en un flujo a lo largo del anillo siguiendo la trayectoria del haz luminoso. Como consecuencia, incluso cuando el anillo no experimenta una rotación, las ondas estacionarias pueden girar.

Los fabricantes de giroscopios de láser anular reducen al mínimo este problema diseñando anillos en los que hay dos flujos gaseosos contrapuestos. Por ejemplo, un giroscopio puede montar un electrodo negativo y dos positivos estando éstos colocados en dos lados opuestos de un cuadrado. Se elimina así en su mayor parte el flujo gaseoso inducido. La cancelación mutua no es completa, pero una deriva residual estable puede medirse y ser compensada, y, en consecuencia, sólo quedan como causa de error de deriva los cambios imprevistos.

En la utilización del instrumento, la segunda y más seria fuente de error reside en un fenómeno denominado anclaje de frecuencia. Se produce porque la onda estacionaria se inmoviliza en el anillo de modo que un observador fijo en el anillo no puede decir si éste gira o no. El anclaje de frecuencia, cuyos efectos sólo se presentan si la velocidad de rotación es relativamente pequeña, desempeña un papel análogo al del rozamiento en un giroscopio mecánico. La onda luminosa estacionaria puede considerarse la superposición de dos ondas que se propagan a lo largo del anillo en direcciones opuestas. El anclaje de frecuencia procede de un acoplamiento interactivo entre las dos ondas luminosas. Este acoplamiento es debido a que las minúsculas imperfecciones de los espejos del giroscopio dan lugar a la dispersión de una pequeña fracción de la onda luminosa incidente en dirección opuesta a la de su trayectoria inicial.

Se puede comprender el efecto resultante del acoplamiento entre las ondas incidentes y las ondas dispersadas que retroceden considerando un sencillo modelo. Supongamos que los espejos sean perfectos y que se fije dentro del anillo, perpendicular a la trayectoria de la luz, una delgada lámina de vidrio. (Ya que las imperfecciones de los espejos están fijas en el anillo también lo deberá estar la lámina de vidrio.) Una pequeña cantidad de la luz que incide en el vidrio es reflejada en dirección opuesta. La onda estacionaria del giroscopio de láser anular tiende a presentar un nodo sobre la lá-



4. ANCLAJE DE FRECUENCIA. Constituye una posible fuente de error en el funcionamiento de un giroscopio de láser anular. Provoca que la onda estacionaria quede "anclada" en el anillo del instrumento, de modo que el observador situado en el anillo no puede decir si el aparato gira o no. La onda estacionaria puede considerarse el resultado de superponer dos ondas que se propagan a lo largo del anillo en sentidos opuestos. El fenómeno del anclaje obedece a minúsculas imperfecciones en los espejos del giroscopio. Tales imperfecciones ocasionan que una pequeña fracción de la onda luminosa incidente se disperse en dirección opuesta a la trayectoria original (*izquierda*). Aparece entre las ondas un acoplamiento cuyas consecuencias pueden comprenderse imaginando que los espejos son perfectos y se ha colocado una delgada lámina de vidrio dentro del anillo, perpendicularmente a la trayectoria seguida por la luz (*derecha*). La onda estacionaria tiende a presentar un nodo ("valle", o punto estacionario) en el lugar ocupado por la lámina de vidrio. Puesto que ésta se encuentra fijada al anillo, la onda estacionaria tiende a girar con el anillo. Esto ocurre cuando las velocidades de rotación son pequeñas, presentándose así el anclaje de frecuencia. El efecto puede reducirse haciendo oscilar el anillo a un ritmo elevado mediante un vibrador; el movimiento "deja suelta" la onda estacionaria.

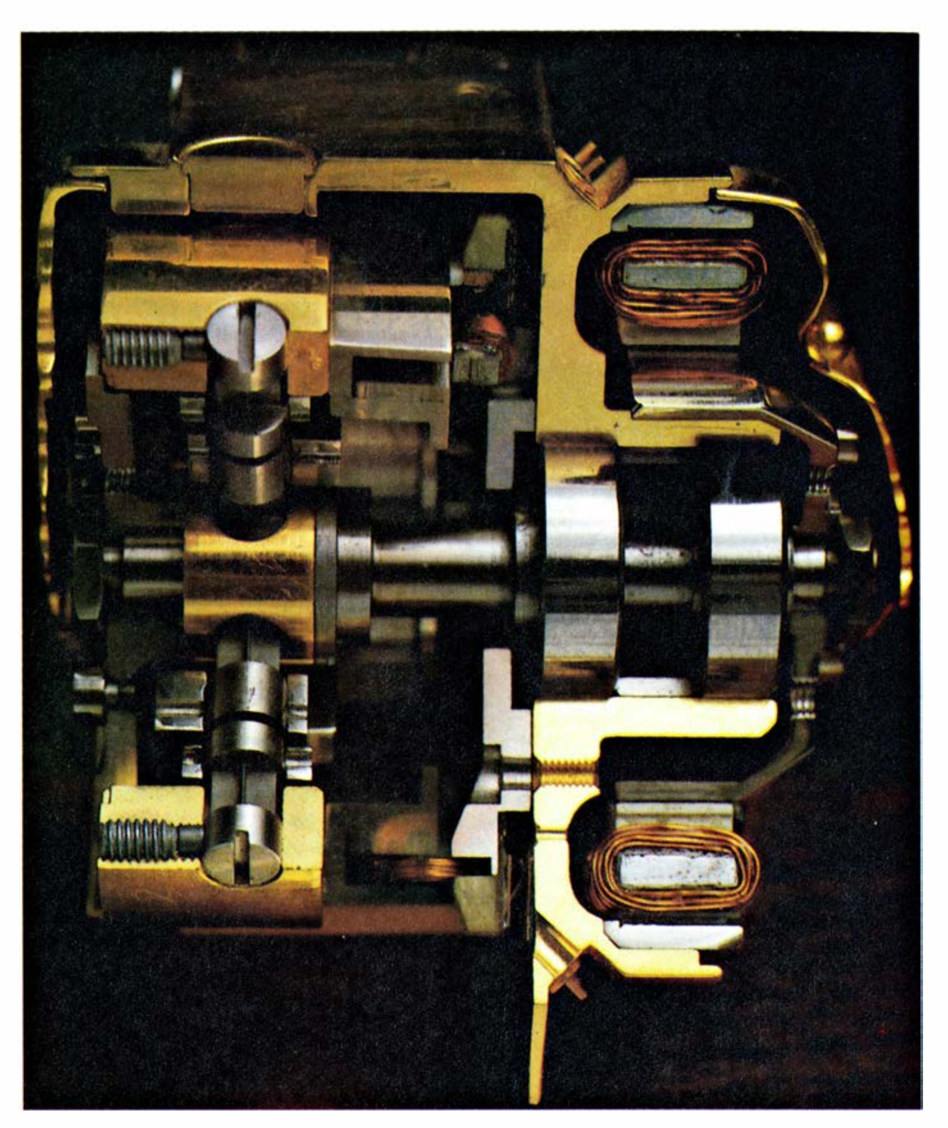
mina de vidrio. Por ser éste fijo, la onda estacionaria tenderá a girar con el anillo, y si la velocidad de rotación es baja, consigue hacerlo.

Una analogía válida del anclaje de frecuencia la constituye una bola sumergida en un líquido viscoso, como un almíbar que descienda por una pendiente ondulada. La fuerza de gravedad representa en este caso la rotación del anillo. Una ladera con fuerte pendiente corresponde a una elevada velocidad de rotación. El fluido viscoso impide que la bola se acelere indefinidamente; impone, pues, un límite a la velocidad. Las ondulaciones representan los nodos y los antinodos de la onda estacionaria y la bola corresponde a la posición de la lámina de vidrio. La altura de las ondulaciones es proporcional a la cantidad de luz que la lámina refleja.

Sin rotación no hay pendiente y la bola permanece en uno de los valles localizados entre dos ondulaciones, o sea, en un nodo. Cuando el anillo gira lentamente, en nuestra analogía sólo debe existir una ligera pendiente. La bola se movería un poco hacia abajo, pero en la siguiente ondulación encontraría una ligera subida que no podría remontar y quedaría atrapada. Esto corresponde al anclaje de frecuencia, es decir, se presenta una situación en la cual la onda estacionaria permanece fija respecto al anillo y el giroscopio no da ninguna señal, con lo que falla la detección de la rotación.

A las velocidades de rotación más altas, les corresponde en nuestro modelo una mayor inclinación y la bola puede rebasar las ondulaciones descendiendo por la pendiente. Tales inclinaciones corresponden a velocidades de rotación para las que las frecuencias quedan desancladas y el anillo ya es sensible a la rotación: la onda estacionaria no permanece fija respecto al anillo y el grisocopio da una señal. Conforme se aumenta la velocidad de rotación, la perturbación en el movimiento de la bola ocasionada por las irregularidades es cada vez menos importante.

El anclaje de frecuencia a bajas velocidades de rotación ha entorpecido la industria de giroscopios ópticos desde sus inicios. Se han dedicado grandes esfuerzos para eliminar el problema. La solución que ha tenido más éxito consiste en el empleo de un vibrador mecánico que hace girar rápidamente el anillo alternativamente en uno y otro sentido. En líneas generales, lo que se pretende es mantener en movimiento el anillo para evitar el anclaje de frecuencia. Puesto que la oscilación del anillo no tiene como resultado ningún giro, la vibración no afecta al resultado de las mediciones. Sin embargo, resulta enojoso tener que acudir a este tipo de solución, puesto que en principio el giroscopio de láser anular presentaba la gran ventaja de carecer de partes móviles. Aunque el movimiento de vibración es suave, el vibrador resulta complicado (de todos modos, mucho menos que el giroscopio mecánico).



5. GIROSCOPIO MECANICO. Aunque funciona de acuerdo con el mismo principio que un giroscopio de juguete, resulta mucho más complejo. Sus aplicaciones en la navegación están pasando al giroscopio óptico.

Los problemas de deriva y de anclaje de frecuencia son de índole técnica y en principio pueden eliminarse de raíz. En último término, la sensibilidad del instrumento puede resultar afectada por causas más determinantes: los dictados de la mecánica cuántica y del principio de incertidumbre de Heisenberg. En su forma más simple el principio de incertidumbre establece que es imposible conocer simultáneamente la posición y la velocidad exactas de una partícula, de un electrón, por ejemplo. Aplicado al giroscopio de láser anular, esto significa que una onda estacionaria no permanecerá rigurosamente en reposo ni siquiera aunque el anillo se mostrara estacionario. Merece destacarse que la tecnología haya avanzado tanto que permita que los giroscopios de láser anular lleguen a alcanzar, dentro de un factor mejor que 10, los límites impuestos por la mecánica cuántica.

La causa fundamental de error en un giroscopio de láser anular reside en una manifestación del principio de incertidumbre conocida con el nombre de emisión espontánea. La energía suministrada excita los electrones de los átomos que constituyen el medio amplificador en el anillo óptico. El fenómeno láser se produce cuando un fotón, o cuanto de luz, estimula un átomo y ocasiona la transición de un electrón excitado a un estado inferior no excitado. Cuando ocurre tal transición, el átomo emite un fotón que tiene la misma dirección, frecuencia y fase que el fotón incidente. Pero es frecuente que un electrón en el átomo experimente una

transición espontáneamente, emitiendo un fotón con una dirección y fase arbitrarias. Algunos de esos fotones son emitidos según la dirección de la onda estimulada y cambian la fase de la onda en una pequeña cantidad aleatoria. En el transcurso del tiempo, las emisiones espontáneas provocan que la onda estacionaria se aparte de su posición inicial: el giroscopio desarrolla un error de corrimiento angular.

En general, para aplicaciones que requieren fina precisión se necesitan giroscopios grandes, pues así se reduce al mínimo el influjo del error de corrimiento. La calidad de los espejos afecta también a la precisión. Un espejo perfecto reflejaría todos los fotones incidentes. En realidad, cada vez que un fotón incide en un espejo existe una pequeña probabilidad de que sea dispersado, absorbido o transmitido. El grado de rendimiento de un giroscopio aumenta con el número de vueltas realizadas por cada fotón antes de que se pierda por absorción, transmisión o dispersión. Por este motivo, el desarrollo de los giroscopios de láser anular ha propiciado un espectacular progreso en la técnica de espejos libres de pérdidas. Un buen espejo de tales giroscopios alcanza un ritmo de pérdidas de fotón inferior al uno por cada 5000 fotones incidentes. (En un espejo común de cuarto de aseo se pierde un fotón de cada 20 incidentes.)

A pesar de las evidentes diferencias entre los giroscopios de fibra y los de láser anular, la naturaleza impone, mediante la mecánica cuántica, que sus rendimientos teóricos resulten similares. Un carrete de fibra con un cierto número de espiras equivale, en el fondo, a un giroscopio de láser anular cuyos fotones recorren el mismo número de vueltas en el anillo, suponiendo que los instrumentos tengan las mismas dimensiones, posean idénticas potencias ópticas y operen con luz de la misma longitud de onda.

Desde luego, ningún giroscopio óptico fascina tanto a los chiquillos como el juguete que constituye la versión simple del giroscopio mecánico. Claro que el éxito de cualquier instrumento no depende de tales consideraciones, sino que va ligado a una cuestión esencialmente económica. El impulso que recibe la tecnología de los giroscopios ópticos está relacionado con el coste de éstos. Por otra parte, la belleza de los principios en que se sustentan los giroscopios ópticos y su elegante estructura es algo más que un mero reflejo de su valor como instrumento para la navegación.

Las tensiones del nacimiento

El esfuerzo propio del recorrido del canal de nacimiento no perjudica a la mayoría de los bebés. De hecho, la afluencia de hormonas "de tensión" que desencadena el parto parece favorecer la supervivencia del recién nacido

Hugo Lagercrantz y Theodore A. Slotkin

A primera vista, el nacimiento no deja de constituir una terrible y peligrosa prueba. El feto humano tarda varias horas en su avance por el estrecho canal de nacimiento, durante las cuales la cabeza soporta una presión considerable y el bebé se ve privado de oxígeno de forma intermitente (debido a la compresión que sufren la placenta y el cordón umbilical en cada contracción uterina). Abandona luego un medio cálido, oscuro y protegido para encontrarse en una fría y luminosa sala de hospital, donde un ser enorme le sostiene cabeza abajo y a menudo le propina unos azotes en el trasero. Además, durante los esfuerzos del nacimiento –especialmente debidos a la hipoxia (privación de oxígeno) y a la presión que se ejerce sobre la cabeza–, el feto suele presentar elevados niveles de hormonas "de tensión", adrenalina y noradrenalina, superiores a los que se registran, por ejemplo, en una mujer que da a luz o en una persona que sufre un ataque cardiaco. La adrenalina y la noradrenalina, representantes principales de la especie química de las catecolaminas, disponen al cuerpo para luchar o huir ante la percepción de cualquier amenaza contra la supervivencia (una tensión); la presencia de elevadas concentraciones de esas hormonas indica, por lo común, que el organismo se encuentra en peligro.

A pesar de las apariencias, las tensiones de un parto normal no suelen resultar perjudiciales. Las pruebas reunidas por nosotros y otros investigadores durante las dos últimas décadas señalan que el feto se encuentra bien dotado para hacer frente a dicha tensión, incluso en las primeras etapas de la gestación; de hecho, son las catecolaminas las que aportan gran parte de la protección en situaciones tan adversas como la hipoxia. Más aún, resulta importante atravesar las circunstancias que provocan la producción de hormonas de tensión. La consiguiente afluencia de hormonas prepara al bebé para sobrevivir fuera del útero materno. Despeja los pulmones y modifica sus características fisiológicas, facilitándole una respiración normal, moviliza combustible de aprovechamiento inmediato para alimento de las células, garantiza el aporte adecuado de sangre al corazón y al cerebro y quizás estimula la vinculación afectiva entre madre e hijo.

El primer indicio importante de la contribución de las catecolaminas a la supervivencia surgió con los estudios pioneros realizados, a finales de la década de 1960, por Robert S. Comline y Marian Silver, de la Universidad de Cambridge. Comprobaron esos autores que, en el feto de vaca, caballo y oveja, las glándulas suprarrenales producían catecolaminas en respuesta a la hipoxia incluso antes de que las inervara el sistema simpático, hallazgo éste muy valioso. Los animales maduros sólo alcanzan elevados niveles de catecolaminas al activarse el sistema nervioso simpático: sus nervios esplácnicos, que proceden del sistema nervioso central, estimulan la glándula suprarrenal para que vierta a la sangre hormonas de tensión. El sistema nervioso simpático, junto con el parasimpático, inerva todos los órganos importantes y contribuye a regular actividades de carácter involuntario, como la respiración, la circulación o el control de la temperatura corporal.

Cuando el sistema nervioso central percibe una amenaza, envía impulsos a los nervios simpáticos para que liberen noradrenalina en los tejidos efectores pertinentes, desencadenándose respuestas rápidas y localizadas de "lucha o huida". Por otra parte, los nervios esplácnicos, que se sirven de la acetilcolina como neurotransmisor, activan la médula suprarrenal (la porción interna de la glándula suprarrenal) instándola a segregar adrenalina y, en menor medida, noradrenalina; pasan esas hormonas a la circulación sanguínea y desencadenan un aumento, suave pero de carácter general y prolongado, del estado de atención. En concreto, las catecolaminas aceleran el ritmo cardiaco, incrementan la fuerza de la contracción del músculo cardiaco y dilatan o constriñen de manera selectiva los vasos sanguíneos, derivando la sangre de aquellos órganos cuya actividad no resulte imprescindible durante la amenaza (piel, intestinos y riñones) y bombeándola hacia aquellos otros de importancia vital (corazón, cerebro y musculatura esquelética). Las hormonas de tensión también dilatan los bronquiolos, para mejorar la respiración, desencadenan la degradación de grasa y glucógeno en combustibles de empleo inmediato, dilatan las pupilas y erizan el pelo.

Los hallazgos de Cambridge dieron pie a la búsqueda del mecanismo en virtud del cual la presencia de catecolaminas alcanza tan elevados niveles en el feto, e igualmente suscitaron numerosos interrogantes. ¿Cuál es el patrón normal de producción de catecolaminas a lo largo de la gestación? ¿En virtud de qué procesos protegen las catecolaminas al feto durante la privación de oxígeno? ¿Difieren los efectos de las catecolaminas entre el feto y el adulto?

Se comprobó que la producción de catecolamina fetal en ausencia de un sistema nervioso simpático competente constituía una respuesta directa a la tensión (independiente de la activación mediante impulsos nerviosos) por parte de la médula suprarrenal. La glándula es bastante grande, en proporción al tamaño que presenta en el adulto, y produce más noradrenalina que adrenalina a lo largo de la mayor parte de la gestación. La noradrenalina y la adrenalina suelen ejercer efectos similares, salvo que la noradrenalina, por ejemplo, no provoca la afluencia de sangre hacia la musculatura esquelética y sí un

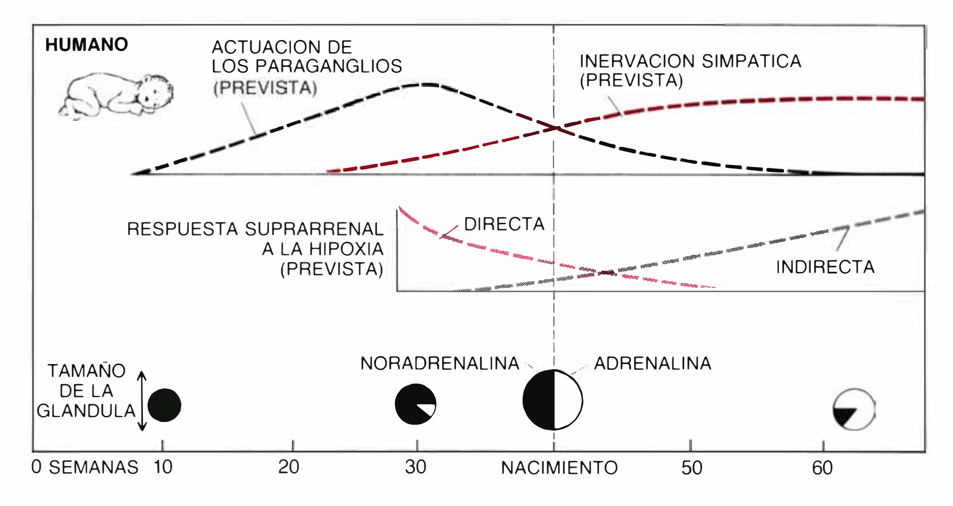
reflejo de deceleración del ritmo cardiaco (mediado por el nervio vago parasimpático) en vez de su incremento. También se descubrió que el feto poseía otras fuentes de noradrenalina: unos tejidos especializados que reciben el nombre colectivo de paraganglios, nódulos cobrizos que salpican la superficie exterior de la aorta y contienen cantidades significativas de noradrenalina. Desaparecen durante la infancia.

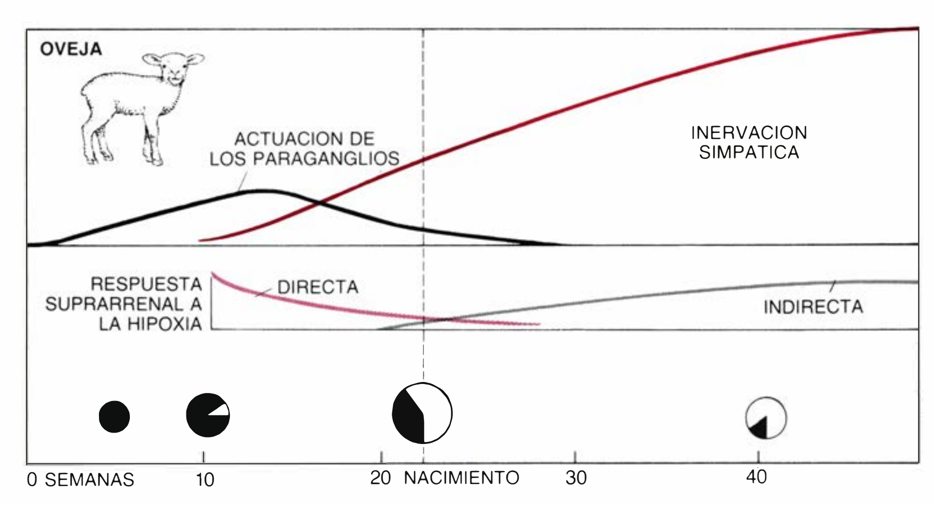
En la Facultad de Medicina de la Universidad de Duke, Frederic J. Seidler, Jorge V. Bartolome y uno de los autores (Slotkin) demostraron que el desarrollo del sistema nervioso, en concreto de los haces que inervan la glándula suprarrenal, provoca que ésta pierda su capacidad secretora de catecolaminas en respuesta directa a la tensión. Experimentaron con ratas recién nacidas, que desarrollan el sistema nervioso simpático después del nacimiento y que, por tanto, fían únicamente en fuentes no nerviosas de catecolaminas, en concreto la médula suprarrenal. En primer lugar, demostraron que las catecolaminas ofrecen a ese animal de sistema nervioso inmaduro una protección relativamente buena contra la hipoxia. En situaciones de baja presencia de oxígeno, esas ratas recién nacidas segregaban grandes cantidades de catecolaminas y lograban sobrevivir aun cuando se prolongara e intensificara la hipoxia. Si se les extirpaban las suprarrenales al nacer, o se les administraba drogas que bloquearan los efectos de las catecolaminas, perdían su capacidad para sobrevivir a la hipoxia.

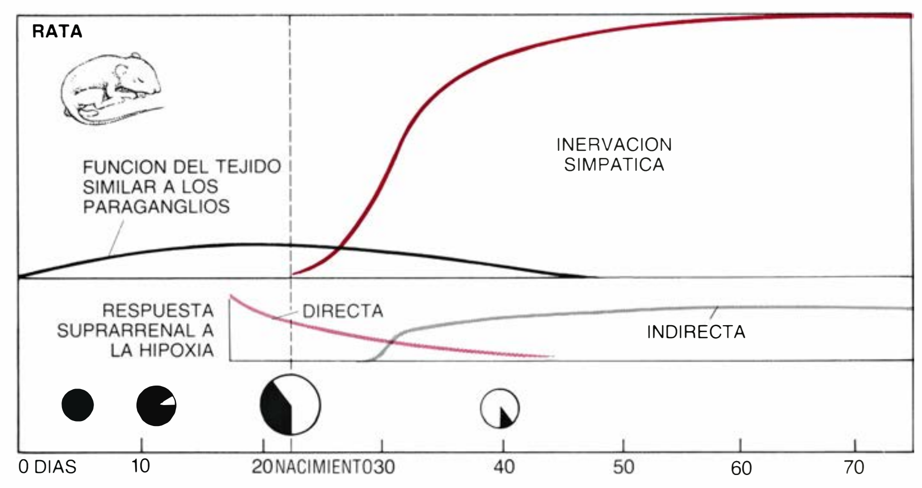
El descubrimiento de que la maduración del sistema nervioso desencadenaba la pérdida de la capacidad de la glándula suprarrenal para responder directamente a la tensión se produjo al forzar de forma prematura la adquisición de competencia por parte de los nervios simpáticos de las ratas. Las ratas manipuladas en ese sentido sólo podían ya liberar catecolaminas por secreción suprarrenal nervo-dependiente, y eran menos resistentes a la hipoxia que las ratas control. Por otra parte, si se anulaba la inervación suprarrenal seccionando la conexión nerviosa antes de que adquiriera compe-



1. RECIEN NACIDO CON SU MADRE inmediatamente después de un parto vaginal normal. El bebé alumbrado por vía vaginal presenta un elevado nivel de las llamadas hormonas de tensión, adrenalina y noradrenalina, que pertenecen a la especie química conocida por catecolaminas. Señalan las pruebas que ese incremento hormonal durante el parto realza la capacidad del neonato para sobrevivir tras el corte del cordón umbilical, y que los que no experimentan el aumento se hallarían en desventaja. También le protege de la asfixia durante los cortes del aporte de oxígeno que se producen en cada contracción uterina.







2. DESARROLLO DE FUENTES DE CATECOLAMINAS en los fetos de rata, oveja y humano. En el organismo maduro (que no se muestra) las principales fuentes de catecolaminas en plasma son los nervios simpáticos que segregan noradrenalina y (en menor medida) las glándulas suprarrenales, que producen más adrenalina que noradrenalina. Las glándulas suprarrenales adultas están sometidas al control de los nervios esplácnicos, fibras simpáticas que transmiten señales por medio de la acetilcolina. En el feto, esos nervios simpáticos no funcionan en los primeros estadios de la gestación. Hasta que adquieren competencia, el bebé en desarrollo se apoya fundamentalmente en las glándulas suprarrenales en lo que se refiere al aporte de catecolaminas. Las glándulas suprarrenales fetales (proporcionalmente mayores que en el adulto y que segregan sobre todo noradrenalina) no dependen de que las activen impulsos nerviosos, sino que liberan directamente las catecolaminas en respuesta a la tensión; más adelante, esa respuesta directa de las suprarrenales desaparece. El feto posee también otro tejido que libera noradrenalina, los denominados paraganglios, de mínima expresión en la rata y muy prominentes en el feto humano. En la rata (*abajo*), que desarrolla la función nerviosa simpática después del nacimiento, la glándula suprarrenal conserva la capacidad de responder directamente a la tensión hasta unos cuantos días después del nacimiento. En la oveja (*centro*), que empieza a desarrollar el sistema nervioso simpático mediada la gestación, la glándula pierde casi toda capacidad de respuesta directa al llegar el nacimiento. Parece que en el hombre (*arriba*) el desarrollo de la inervación simpática progresa más lentamente que en la oveja, y que la glándula suprarrenal conserva su capacidad de respuesta directa a la tensión a lo largo de algunas semanas después del nacimiento del bebé.

tencia, persistían el mecanismo fetal de respuesta directa y su efecto protector sobre la supervivencia.

En las ratas, donde la inervación simpática se desarrolla inmediatamente después del nacimiento, puede demostrarse que la médula suprarrenal responde directamente a la tensión hasta unos pocos días después del alumbramiento. En ovejas, donde la inervación comienza mediada la gestación y al nacer ya ha concluido casi, la capacidad de la glándula suprarrenal para responder directamente a la hipoxia ha desaparecido virtualmente en el momento del alumbramiento. La madurez del recién nacido humano es inferior a la de la oveja, pero superior a la de la rata; quizá pueda su glándula suprarrenal responder directamente a la hipoxia en el nacimiento. En cualquier caso, desconocemos el calendario exacto de maduración del sistema nervioso simpático, así como el del relevo del sistema suprarrenal.

Dos fueron los caminos seguidos a la hora de determinar el patrón natural de liberación de catecolaminas durante la gestación. Uno de nosotros (Lagercrantz), en un trabajo realizado con colegas del Instituto Carolino de Estocolmo, midió la concentración de metabolitos (productos de degradación) de la adrenalina y la noradrenalina en el líquido amniótico que rodea al feto humano. Los niveles de metabolitos en situación de reposo eran bastante bajos, si bien aumentaban a medida que se aproximaba el término de la gestación y se elevaban claramente en situaciones de dolor o angustia, por ejemplo cuando una toxemia materna retardaba peligrosamente el crecimiento fetal en útero.

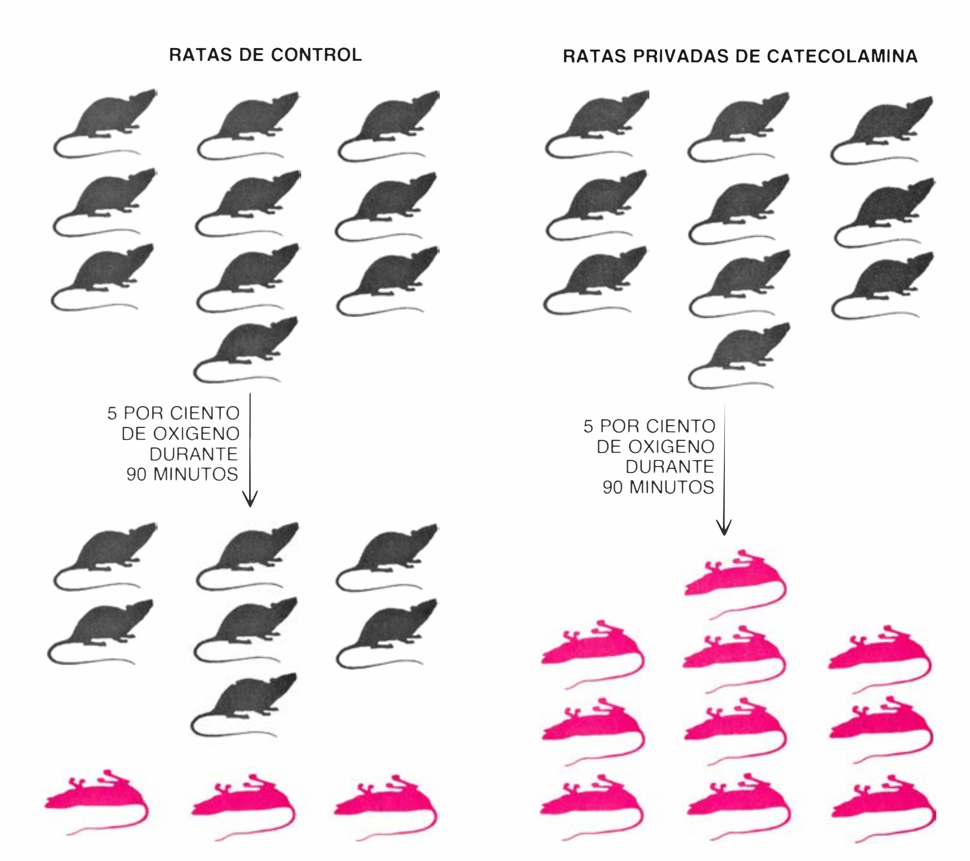
Sin duda resulta útil esa medición indirecta, pero la evaluación definitiva de la respuesta fetal de las catecolaminas exige un control directo. Colin Jones y su grupo, de la Universidad de Oxford, decidieron por ello investigar la liberación de catecolaminas en un modelo animal. Sin desprenderle la placenta, extrajeron un feto de oveja del útero de su madre, le conectaron catéteres y electrodos y lo devolvieron al mismo útero. Jones comprobó que las concentraciones de catecolamina en el feto de oveja en reposo eran muy bajas hasta pocos días antes del nacimiento. Se centuplicaban esas concentraciones en los episodios de asfixia: las hormonas procedían fundamentalmente, de la médula suprarrenal.

Jones, entre otros, describió el proceso en virtud del cual las catecolami-

nas protegen de la asfixia al feto. Cuando se le administró al feto de oveja una inyección de noradrenalina, de concentración comparable a la desencadenada por un episodio de hipoxia, se comprobó la orientación del flujo sanguíneo hacia el corazón, cerebro, suprarrenales y placenta, así como su derivación de otros órganos que no resultan vitales en esas circunstancias. Tales alteraciones hemodinámicas provocaban la elevación de la presión sanguínea y, a su vez, una deceleración refleja inmediata del ritmo cardiaco. (En esa deceleración del corazón interviene el nervio vago parasimpático, que adquiere la funcionalidad antes que los nervios simpáticos.) La reducción del ritmo cardiaco favorece la supervivencia, pues reduce el trabajo del corazón y su necesidad de oxígeno. K. G. Rosén e Ingemar Kjellmer, de la Universidad de Gotemburgo, así como nuestro grupo de Duke, demostraron posteriormente que cualquier interferencia de los efectos cardiovasculares instados por las catecolaminas evita que se den esos cambios adaptativos en situaciones de hipoxia, lo que provoca el colapso circulatorio y la muerte. (Los efectos de las catecolaminas sobre el feto privado de oxígeno mimetizan el reflejo de buceo que se observa en animales marinos. Cuando introducen el rostro en el agua, el flujo de sangre se restringe inmediatamente a los órganos vitales: el consiguiente incremento de la presión sanguínea desencadena el descenso del ritmo cardiaco.)

La respuesta del feto a la tensión es más limitada que la del adulto. Este último, capaz de luchar o huir en respuesta a la tensión, está dotado para incrementar el ritmo cardiaco y enviar sangre rica en oxígeno a los músculos en respuesta a dicha tensión. Por el contrario, el feto debe soportar la privación de oxígeno: aunque exquisitamente preparado para ello, no lo está para mucho más. De hecho, el feto está mejor dotado para soportar una situación de asfixia que el adulto, que por regla general presenta irregularidades cardiacas a los pocos minutos.

¿Por qué reducen las catecolaminas el ritmo cardiaco en el feto y no en el adulto? Quizá porque éste, a diferencia de aquél, produce gran cantidad de adrenalina en respuesta a la tensión, mientras que el feto libera sobre todo noradrenalina: el incremento de los niveles de adrenalina puede compensar una deceleración parasimpática del corazón, lo que no ocurre en el caso de la noradrenalina. (Llegado el momento del parto, el bebé sí libera adrenalina suficiente para que se den incrementos o decrementos del ritmo cardiaco.) Una segunda razón: como ha demostrado uno de los autores (Slotkin), el corazón inmaduro es menos sensible; para desencadenar el aumento de ritmo cardiaco se exige una concentración superior de adrenalina, de ahí que resulte menos probable que la afluencia de catecolaminas provoque alteraciones del latido.

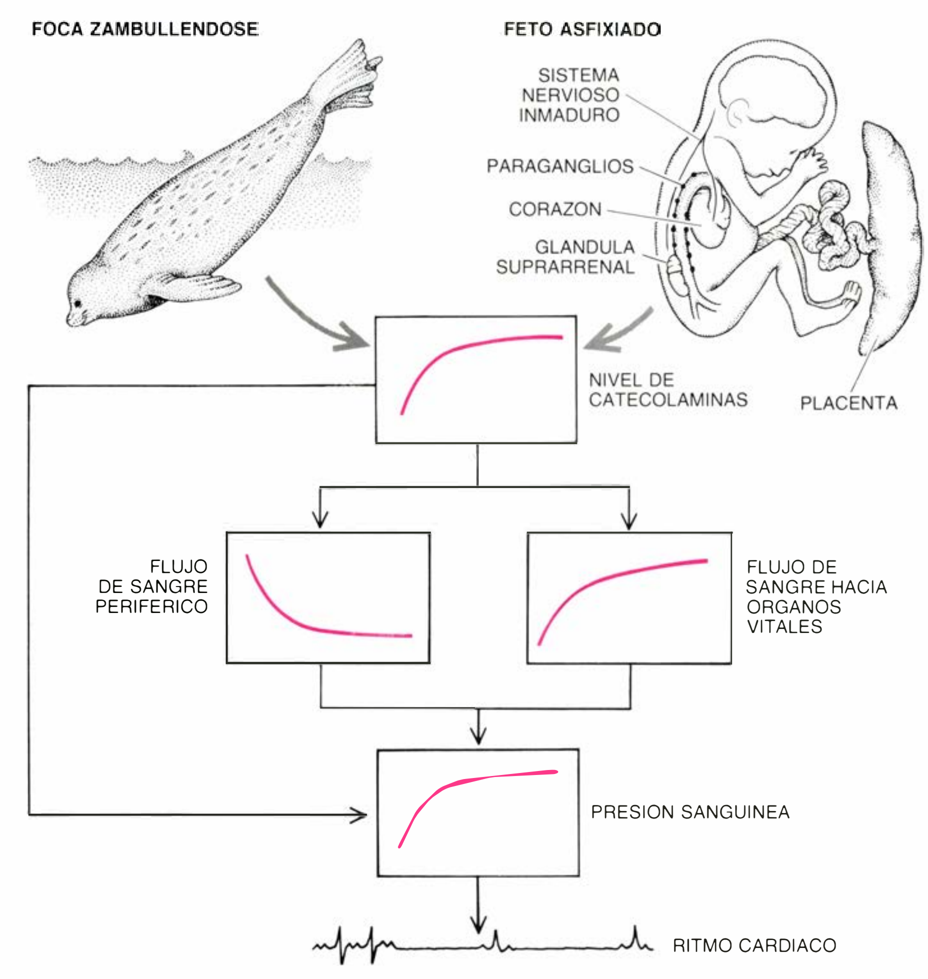


3. PRIMERAS PRUEBAS que demostraron la necesidad del incremento de catecolaminas para que el feto o el recién nacido sobrevivan a la hipoxia (privación de oxígeno). Cuando se extirparon las glándulas suprarrenales a ratas de un día de edad (que, como el feto humano, poseen un sistema nervioso todavía inmaduro), o se les administró drogas que bloquean el efecto de las catecolaminas, ninguna sobrevivió a 90 minutos de escasez de oxígeno, mientras que el 70 % de las ratas de control sobrevivieron al mismo ensayo.

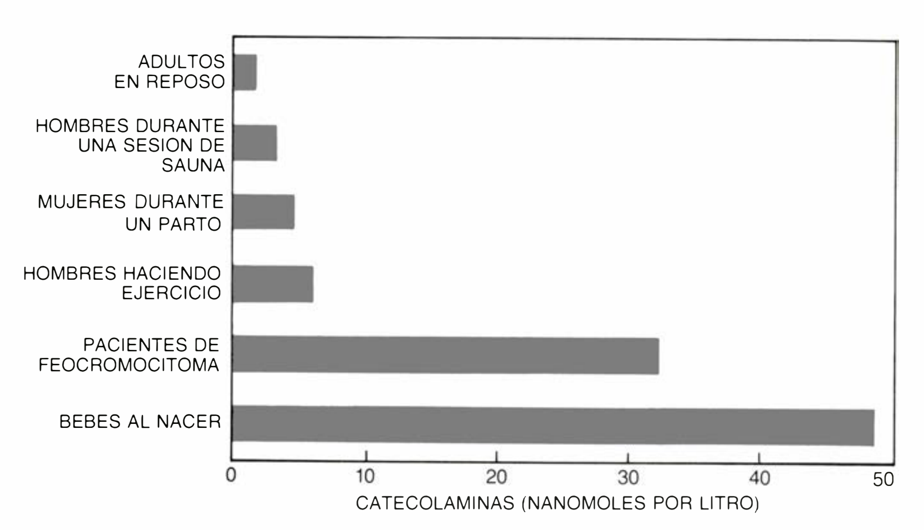
El descubrimiento de que los niveles de catecolaminas se mantienen bajos a lo largo de la mayor parte de la gestación salvo en situaciones de amenaza vino a sugerir que, como en el adulto, el pico de catecolaminas actúa en el feto a modo de sistema de protección, muy eficaz en casos de emergencia. Tras investigaciones realizadas sobre el posible abuso de la extracción por cesárea nació la sospecha de que las hormonas no sólo protegían al feto durante la aguda tensión que sufría en los períodos de asfixia, sino que también preparaban al cuerpo para su actuación fuera del útero.

Hace algunos años, la proporción de partos con cesárea en Estados Unidos y Europa había aumentado, desde el 5 por ciento, hasta el 15 o 20 por ciento. Y ello porque las modernas técnicas de control permitían detectar cualquier alteración del ritmo cardiaco fetal en el transcurso de las contracciones uterinas; cuando se advertían cambios complejos durante las contracciones (no ya simples incrementos o deceleraciones del ritmo cardiaco), solía considerarse que el bebé se asfixiaba y su vida corría peligro, procediéndose a extraerlo por medios quirúrgicos. Sin embargo, al nacer, se comprobaba que más del 50 por ciento de los bebés no presentaba signo clínico alguno de asfixia.

Juzgando que quizá muchas de las cesáreas fueran innecesarias, uno de nosotros (Lagercrantz), junto con Peter Bistoletti y Lars Nylund, del Instituto Carolino, decidió averiguar si la secreción normal de catecolamina en respuesta al esfuerzo materno explicaba la aparición de latidos complejos advertida en muchos fetos. Se comprobó que, en efecto, la afluencia de catecolaminas provocada por el proceso normal de parto alteraba el ritmo cardiaco, alteración que, equivocadamente, podía interpretarse como una señal de sufrimiento fetal a menos que el control se complementara con pruebas bioquímicas. Sólo si el *p*H sanguí-



4. RESPUESTA FETAL A LA ASFIXIA; se asemeja a la de algunos animales marinos cuando se zambullen. Al sumergir la cabeza una foca, aumentan sus niveles de catecolaminas: los nervios simpáticos y las glándulas suprarrenales segregan las hormonas, que provocan un descenso del flujo de sangre periférica y el incremento del flujo de sangre que riega el cerebro, corazón y glándulas suprarrenales. El efecto conjunto de las propias catecolaminas y de los cambios que ellas mismas provocan en el flujo sanguíneo se traduce en un aumento de la presión sanguínea y una deceleración del ritmo cardiaco. De modo similar, cuando un feto humano se ve privado de oxígeno, sus glándulas suprarrenales y paraganglios (así como, en menor medida, los nervios simpáticos, que se encuentran en proceso de desarrollo) desencadenan un aumento de los niveles plasmáticos de catecolaminas. Disminuye entonces el flujo de sangre hacia la periferia y aumenta el que riega el cerebro, corazón, glándulas suprarrenales y placenta; se eleva la presión sanguínea y se frena el ritmo cardiaco.



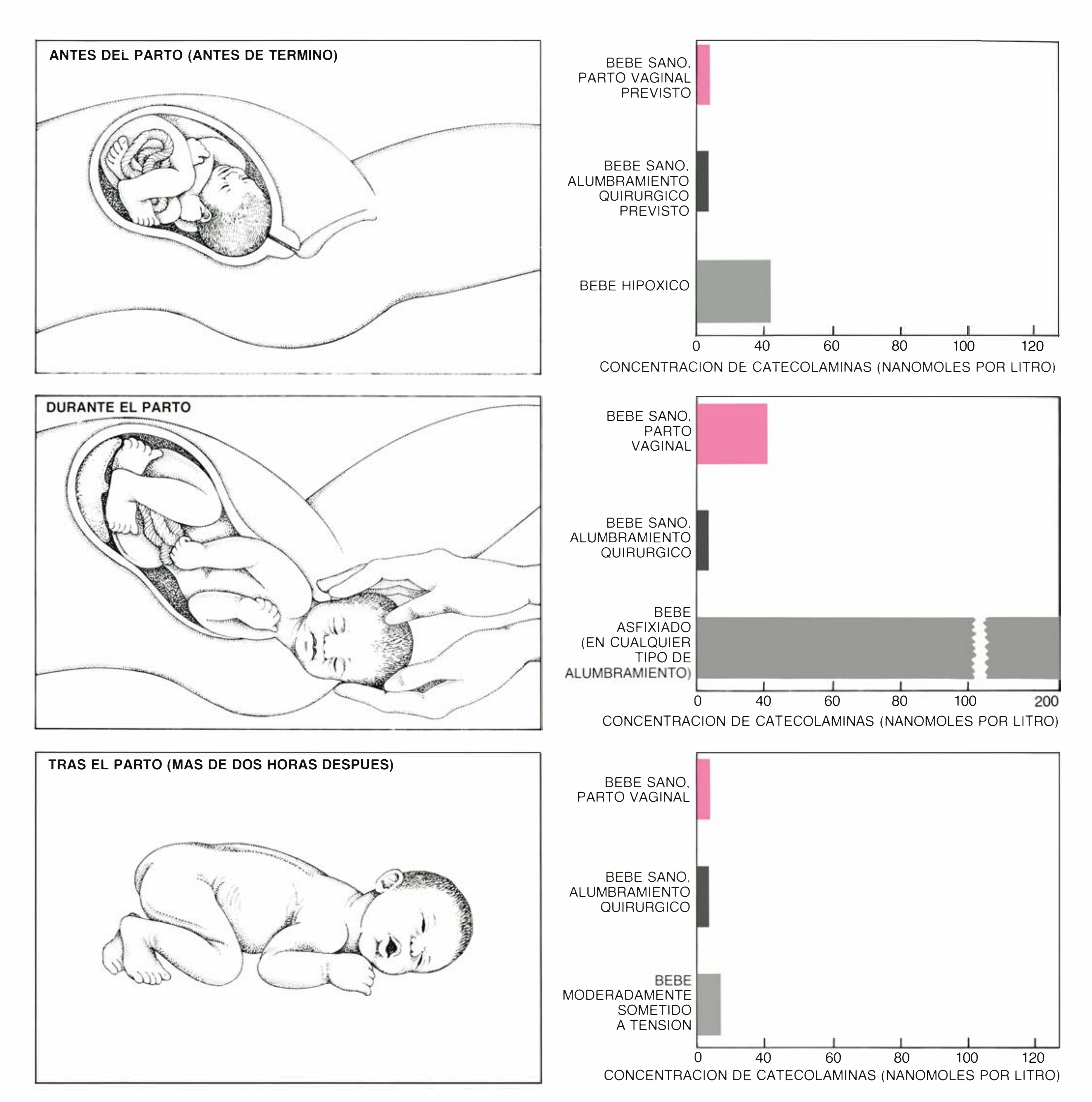
5. NIVELES DE CATECOLAMINAS en adultos y recién nacidos. Las concentraciones registradas en muestras de cordón umbilical de neonato superaban unas 20 veces el valor medio de la sangre venosa en adultos en reposo. También presentaban los neonatos niveles de catecolaminas superiores a diversos adultos en situación de tensión. (Un feocromocitoma es un tumor que segrega catecolaminas.) El hallazgo de que los bebés sanos alumbrados por vía vaginal presentaban niveles catecolamínicos superiores a los propios de adultos sometidos a grave tensión vino a proporcionar uno de los primeros indicios de que el incremento de catecolaminas durante el parto posee una función que trasciende a la de protección del feto frente a la hipoxia.

neo del pericráneo fetal desciende por debajo de 7,25 son señal de asfixia esos aparentes trastornos cardiacos.

Llevamos a la práctica esos controles y comprobamos que, en el bebé humano, el proceso normal del parto desencadena un incremento sorprendentemente elevado de catecolaminas en plasma. Raul Artal, de la Universidad del Sur de California, ha publicado hallazgos similares. Ya al principio del parto normal, cuando el cuello de la matriz de la madre apenas se había dilatado (dos o tres centímetros), las concentraciones de catecolaminas en muestras de pericráneo fetal eran unas cinco veces más altas que en un adulto normal en reposo. Después del nacimiento los niveles de catecolaminas se duplicaban o triplicaban, señal de que habían aumentado a lo largo de la fase siguiente, durante los pujos de expulsión. (Los niveles de catecolaminas declinan por lo general unos 30 minutos después del parto, retornando a niveles de reposo aproximadamente dos horas después del nacimiento.)

En estudios posteriores se puso de manifiesto que los bebés que sí se asfixiaban (*p*H inferior a 7,25), por presentarse de nalgas o estrangulados por el cordón umbilical, registraban niveles de catecolaminas muy elevados, entre 500 y 1000 nanomoles por litro de plasma, niveles que provocarían colapso en un adulto. Tales concentraciones sólo podían alcanzarse si la asfixia provocaba a su vez una mayor liberación de catecolaminas por parte de la médula suprarrenal, paraganglios y nervios.

Parte del incremento que suele registrarse en el parto probablemente se deba a cierto grado de hipoxia; casi todos los recién nacidos presentan un déficit de oxígeno semejante al de un corredor cuando acaba la carrera. Sin embargo, ese déficit de oxígeno no explica por sí solo la elevación habitual de los niveles de catecolaminas; únicamente cuando el bebé se encuentra considebastante asfixiado se comprueba la existencia de una relación directa entre la asfixia (medida a través del *p*H del plasma) y los niveles de catecolamina. Estudios realizados en animales demuestran que la presión sufrida por la cabeza durante las contracciones uterinas puede estimular la actividad simpato-suprarrenal, lo que sugiere que la compresión de la cabeza del feto humano explica en gran parte el incremento de secreción de catecolaminas que se produce durante el esfuerzo. Datos reunidos por Lars Ires-



6. ESTIMACION DE LOS NIVELES de hormonas de tensión en el bebé humano antes, durante y después del parto. (No puede efectuarse en el feto humano una medición directa de las catecolaminas.) Suelen mantenerse bajos los niveles hasta unos pocos días antes del nacimiento (*arriba*), elevándose sólo si el bebé sufre alguna tensión. Se incrementa la concentración durante el parto vaginal (*centro*), en gran parte en respuesta a las contracciones uterinas de la madre (que provocan una privación intermitente de oxígeno) así como a las presiones que sufre la cabeza a medida que el feto avanza por el canal de parto. El pico hormonal protege al bebé contra la asfixia durante el parto, a la vez que facilita la estabilización de sus funciones tras el nacimiento. Si se practica una cesárea antes de que la madre se ponga de parto, el bebé probablemente no llegue a experimentar el incremento de catecolaminas durante el alumbramiento. En el caso de que el bebé sufra graves tensiones en el momento de nacer, como ocurre cuando se le enrolla el cordón umbilical alrededor del cuello, pueden alcanzarse niveles de catecolaminas extraordinariamente elevados; si a esos niveles se les suma el descenso del *p*H está indicada la práctica de un alumbramiento quirúrgico de urgencia. Superadas las primeras horas de vida (*abajo*) las catecolaminas suelen retornar a los niveles de reposo. (Dibujos de Charles R. Belinky.)

tedt, Paul Hjemdahl y uno de los autores (Lagercrantz) respaldan la sospecha de que los bebés alumbrados por cesárea voluntaria y antes de iniciarse el parto presentan bajos niveles de catecolamina, mientras que los alumbrados por cirugía de urgencia (ya desencadenado el parto) presentan un incremento sólo ligeramente inferior al de los bebés alumbrados de parto normal por vía vaginal.

No podíamos ignorar la elevación de las catecolaminas, sorprendentemente acusada, observada en el feto incluso en un parto rutinario. Llevaba a concluir que, además de proteger al feto durante el notable esfuerzo del alumbramiento, el pico de catecolaminas probablemente desempeñara otra función: la de fortalecer la capacidad del neonato para desenvolverse con eficacia desde el instante mismo de separarse de la madre. Trabajos de investigación realizados durante estos últimos años han puesto de relieve que el incremento de catecolaminas facilita la respiración normal y ejerce otros efec-

tos que preparan al bebé para superar la falta de alimento, la privación de oxígeno u otra adversidad que se presente durante las primeras horas de vida.

El fomento de la respiración constituye un efecto adaptativo extremadamente importante del incremento de catecolaminas próximo al nacimiento. Se sabe que los bebés alumbrados por cesárea voluntaria muestran cierta predisposición a padecer trastornos respiratorios. Dos factores principales colaboran en ello: la inadecuada absorción de líquido pulmonar al nacer, que da a los pulmones un aspecto mojado, y la escasa producción de agente tensoactivo; sustancia jabonosa que reduce la tensión superficial en los alvéolos pulmonares (donde se efectúa el intercambio de oxígeno y dióxido de carbono) y los mantiene abiertos. La absorción de líquido pulmonar y la liberación de suficiente agente tensoactivo parecen depender del incremento sostenido de catecolaminas en plasma durante las horas previas al nacimiento.

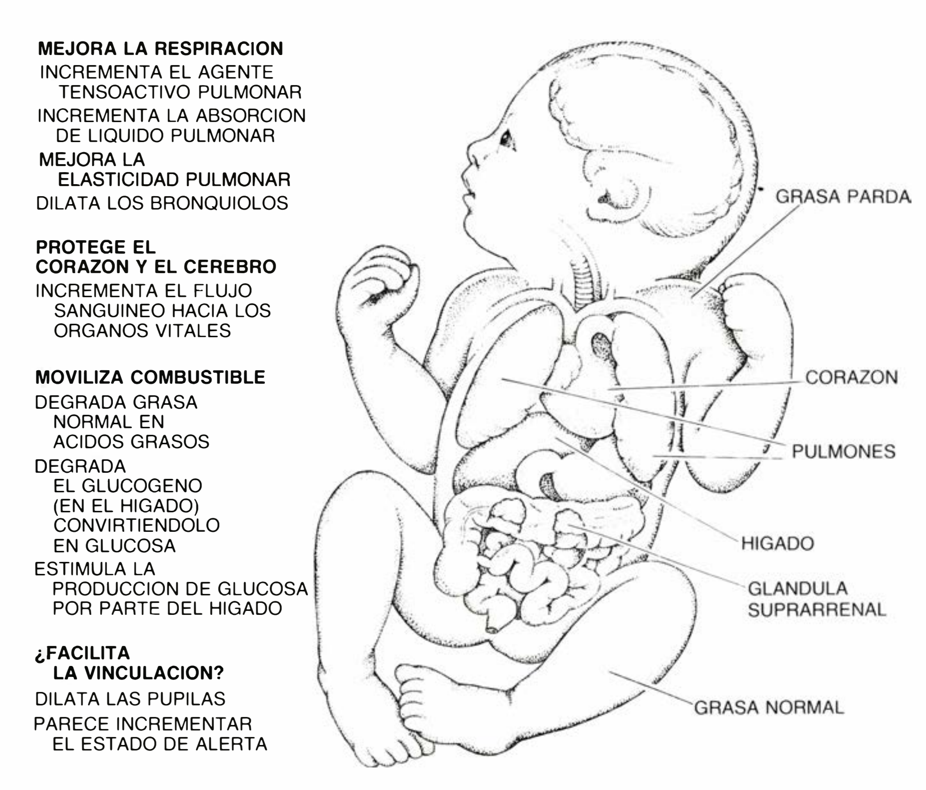
La importancia que reviste el aumento persistente de catecolaminas (especialmente adrenalina) para la absorción de líquido pulmonar en el momento del nacimiento ha quedado demostrada en unos elegantes experimentos realizados con fetos de oveja por Richard E. Olver, Leonard B. Strang y D. V. Walters, del University College de Londres. Se segrega líquido de forma natural en los pulmones del feto de oveja a un ritmo aproximado de 200 a 300 mililitros cada 24 horas; se mueve el líquido por los pulmones como si fuera aire. Ese fenómeno parece de gran interés para el desarrollo alveolar durante la gestación; ahora bien, como sucede en el hombre, debe invertirse el proceso al llegar el nacimiento. Cuando los investigadores inyectaban adrenalina a la oveja, la secreción de líquido pulmonar se detenía inmediatamente y se absorbía dicho líquido. La adrenalina inyectada ejercía su efecto máximo a llegar la gestación a término, por lo que es de esperar que la aparición de catecolaminas al aproximarse el parto resulte importante para el desarrollo de la competencia respiratoria. La medición de adrenalina en plasma y de absorción de líquido pulmonar en el feto de oveja durante el parto confirmaron que la formación de líquido se detiene y sustituye por la absorción a medida que crecen los niveles de adrenalina en las horas que preceden al nacimiento.

Los estudios realizados con ratas, en Duke, confirman el interés vital que el incremento de catecolaminas durante el parto reviste para la función pulmonar y para la capacidad de soportar períodos de escasez de oxígeno durante el alumbramiento y nada más nacer. La administración a ratas de pocos días de drogas obstructoras de los beta-2-receptores, que median los efectos de la adrenalina sobre la producción de agente tensoactivo y la absorción de líquido pulmonar, puso en graves aprietos su capacidad para sobrevivir en condiciones de escasez de oxígeno. Al administrarse esas mismas drogas a animales algo mayores (de ocho días de edad), no se observaron tales dificultades. Aparentemente, el aumento de catecolaminas no sólo resulta esencial para la función pulmonar tras el nacimiento, sino que debe producirse también un par de días antes o después del parto para que mejore la respiración del neonato.

Aunque no puede evaluarse directamente la producción y absorción de líquido pulmonar en los bebés, nuestro grupo del Instituto Carolino (Irestedt, Gerd Faxelius, Kerstin Hagnevik, Bo Lundell, Ingrid Dahlin y Lagercrantz) ha comprobado que los resultados derivados de estudios realizados con animales pueden aplicarse también a los infantes. Medimos la elasticidad pulmonar (la capacidad del órgano para ensancharse y llenarse de aire), que depende en parte de la absorción de líquido pulmonar, y hallamos pocas diferencias entre bebés alumbrados por vía vaginal o quirúrgica inmediatamente después de nacer; dos horas después, en cambio, alumbrados por vía vaginal presentaban una elasticidad pulmonar significativamente mejor. Además, se advirtió una estrecha correlación entre la concentración de catecolaminas al nacer y la elasticidad pulmonar a las dos horas en el grupo alumbrado vaginalmente.

Que no se produjeran cambios de la elasticidad pulmonar inmediatamente después del parto pero sí después indica que esos efectos debíanse a las catecolaminas circulantes; no obedecían a ninguna presión mecánica del tórax durante el parto, que hubiera producido efectos inmediatos. La conclusión de que las catecolaminas facilitan la respiración al nacer la confirma también la observación de bebés nacidos de madres a las que se ha administrado drogas beta-adrenérgicas para evitar el parto antes de término. Esos recién nacidos rara vez sufren arritmias respiratorias graves.

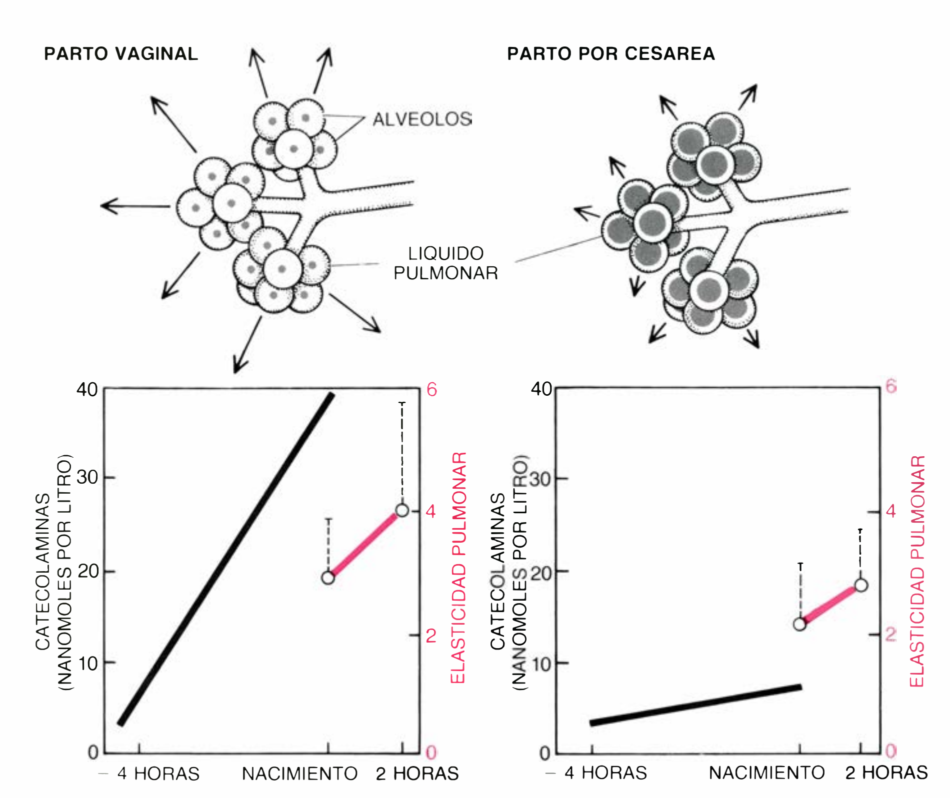
Además de fomentar la respiración normal, el incremento de catecolaminas antes del nacimiento acelera también el ritmo metabólico del bebé du-



7. EFECTOS ADAPTATIVOS del incremento de catecolaminas durante el parto. Contemplan la facilitación de la respiración normal, la alteración del flujo sanguíneo para proteger el corazón y el cerebro ante una posible asfixia, la movilización inmediata de combustible para el aporte de energía y, posiblemente, el refuerzo del vínculo materno-filial. En general, esos efectos preparan al organismo para que se conserve la homeostasis (el funcionamiento estable) durante el parto pese a que quede expuesto a adversidades como la privación de oxígeno. Después del alumbramiento, las catecolaminas activan, en respuesta al frío, un tejido especializado en la producción de calor, la grasa parda, que sólo se encuentra en bebés y niños de corta edad.

rante el parto. Lo que a su vez acelera la degradación de energía almacenada a formas que puedan alimentar a las células en cuanto el bebé deje de recibir el aporte constante de elementos nutritivos por medio del cordón umbilical. Mark A. Sperling, de la Facultad de Medicina de la Universidad de Cincinnati, ha demostrado, en ovejas, que inmediatamente después del parto normal empieza a formarse glucosa, un azúcar. De modo similar, nuestro grupo del Instituto Carolino ha hallado niveles más altos de glucosa, ácidos grasos libres y glicerol (todos ellos productos de la descomposición del combustible almacenado) en la sangre de bebés alumbrados vaginalmente que en la de los nacidos por cesárea voluntaria. También hemos comprobado mayor incidencia de niveles bajos de azúcar en sangre en bebés de parto quirúrgico, señal de que en ellos se registra un déficit en el aporte de la energía preferida por el cuerpo.

El tercer efecto adaptativo principal del pico de catecolamina durante el parto es la alteración del flujo sanguíneo. Nuestro grupo del Instituto Carolino, en colaboración con Alice Yao, del Hospital Clínico Downstate de la Universidad estatal de Nueva York en Brooklyn, ha puesto de relieve que el bebé de parto vaginal presenta al nacer un aumento del flujo sanguíneo hacia los órganos vitales y una reducción de la circulación periférica. Comprobamos que el flujo sanguíneo era inferior en las piernas y superior en los pulmones en el grupo de parto vaginal durante las dos primeras horas de vida, y guardaba correlación con los niveles de noradrenalina en plasma al nacer. Que al nacer se incremente el flujo sanguíneo hacia los órganos vitales quizá no resulte decisivo para todo neonato, pero sin duda eleva la probabilidad de supervivencia del pequeño que experimente dificultades respiratorias.



8. ELASTICIDAD PULMONAR tras el parto por vía vaginal y tras el alumbramiento por cesárea; se muestran además las concentraciones medias de catecolaminas antes del nacimiento, según los datos obtenidos en 27 bebés. La elasticidad (capacidad del pulmón para ensancharse e intercambiar de manera eficaz dióxido de carbono y oxígeno con los capilares a través de los alvéolos) la da la relación entre volumen y presión pulmonares. Las concentraciones de catecolaminas antes del nacimiento se midieron en muestras de sangre pericraneal. Se ha demostrado que los bebés alumbrados por vía vaginal presentan más elasticidad pulmonar.

La revisión de las puntuaciones obtenidas en el test de Apgar, que mide el grado de asfixia de los recién nacidos, corrobora la importancia del aumento de catecolaminas durante el parto en lo que atañe a la capacidad del bebé para sobrevivir a la asfixia. Se desprende de esa revisión que, en general, cuanto mayor es el pico de catecolaminas, tanto mayor es la probabilidad de que el infante supere la privación de oxígeno. El test, que se efectúa de forma rutinaria a la mayoría de los occidentales al nacer, refleja la combinación de cinco valoraciones. La puntuación más alta, 10, se otorga cuando el ritmo cardiaco es fuerte, la respiración regular y el tono muscular adecuado, el bebé reacciona bien a la estimulación externa y el color de la piel es bueno. Aquellos cuya puntuación no llega a 7 probablemente presenten un ritmo cardiaco más lento, hipotonía (escaso tono muscular), cianosis (coloración azul debida a la falta de oxígeno en la sangre) o cualquier combinación similar de trastornos.

Creímos en cierto momento que la concentración de catecolaminas en recién nacidos y la puntuación Apgar guardaban una relación inversa, pues los bebés asfixiados suelen presentar concentraciones de catecolamina superiores a las de los sanos. Investigaciones recientes han demostrado que tal relación es mucho más compleja. Se comprobó que los bebés que sufrían una asfixia moderada durante el parto (de acuerdo con el *p*H en cordón umbilical) alcanzaban puntuaciones Apgar normales al nacer cuando mostraban elevados niveles de catecolaminas, y bajas cuando los niveles de catecolaminas lo eran también. En otras palabras, la secreción de catecolaminas parecía contrarrestar los efectos de la privación de oxígeno.

Más hipotética que el descubrimiento de que las catecolaminas mejoran la adaptación fisiológica a la vida extrauterina es la posibilidad de que el pico de catecolaminas al nacer facilite la vinculación entre madre e hijo. Marshall H. Klaus y John H. Kennel, de la Facultad de Medicina de la Universidad Case Western Reserve, han propuesto que la primera hora que sigue al nacimiento constituye un período decisivo para el vínculo maternofilial; cuando tal vinculación no tiene lugar, la relación entre madre e hijo puede verse perturbada durante algún tiempo. Estos asertos resultan extremadamente polémicos. De ser ciertos, implicarían que una actitud de alerta y dispuesta para la interacción resulta más adaptativa que la indolente y poco atenta.

Sospechamos, de acuerdo con observaciones informales, que el pico de catecolaminas al nacer podría desencadenar esa atención por parte del bebé. Una de las razones que nos llevan a creerlo es que las concentraciones de catecolaminas registradas en el neonato de parto normal durante las primeras horas de vida provocarían en el adulto una actitud de alerta y atención, e incluso en ocasiones hasta les conferiría cierta sensación de bienestar.

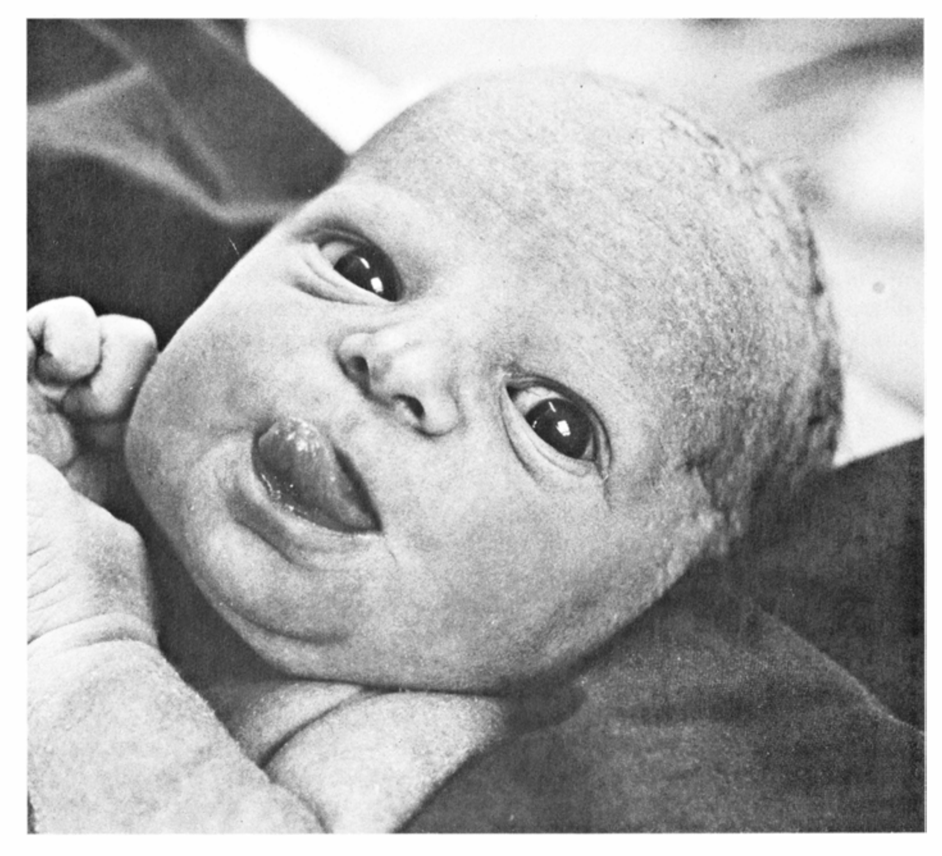
Más allá de sus funciones adaptativas, las catecolaminas son importantes para la supervivencia del recién nacido durante las situaciones de tensión, de igual modo que lo son para el feto o el adulto. Si se priva al bebé de alimento durante varios días (como sucede en al-

gunas culturas, donde no se alimenta al neonato hasta que la madre da leche), la respuesta catecolamínica movilizará el combustible necesario a partir de las reservas de glucógeno y de grasa del organismo. De abandonarse un recién nacido a la intemperie, la respuesta de los terminales simpáticos que inervan la grasa parda, un tejido singular, instará la rápida conversión en calor de los enlaces químicos de sus reservas. La grasa parda, que se concentra debajo de la piel de los hombros y cerca de los riñones en los recién nacidos, disminuye después de la primera infancia. El neonato es particularmente susceptible a padecer rápidas pérdidas de calor, pues su proporción superficie/volumen es muy elevada.

Las pruebas señalan, en conjunto, que el incremento de hormonas de "tensión" en el recién nacido de parto normal no sólo refleja una respuesta a la acusada tensión, sino también un intento, por parte del cuerpo, de aumentar su probabilidad de supervivencia. Sugieren esos hallazgos que los bebés alumbrados por cesárea antes de que la madre se ponga de parto quizá se encuentren en cierta desventaja.

Podría obviarse esa desventaja administrando drogas del tipo de las catecolaminas a los recién nacidos de parto quirúrgico, pero debe antes demostrarse que el método es seguro y eficaz. Por lo pronto, algunos tocólogos, especialmente en Occidente, procuran que el bebé goce de los beneficios del incremento de catecolaminas demorando, si pueden, el alumbramiento quirúrgico hasta que la madre ha superado al menos las primeras fases del parto. Por supuesto, deben evitarse durante el embarazo, y en especial cuando se aproxima su término, las drogas que interfieren la acción de las catecolaminas (como ciertos reductores de la tensión).

Sin duda reconfortará a los padres que, desde la perspectiva del bebé, los esfuerzos del parto normal probablemente resulten menos desagradables y más beneficiosos de lo que pueda sugerir el sentido común. Las fotografías encargadas para un estudio por el desaparecido John Lind, del Instituto Carolino, ilustran ese punto. Lind comprobó que la mayoría de los bebés de la serie presentaban poco después del parto una actitud atenta y jovial, no precisamente infeliz.



9. MUESTRA EL BEBE una actitud alerta a los cinco o seis minutos después de nacer; y también mantiene las pupilas dilatadas pese a hallarse bajo una intensa iluminación, efectos que parecen responder al pico de catecolaminas que se registra durante el parto vaginal. Al provocarle al bebé esa actitud quizás el incremento hormonal facilite la vinculación entre madre e hijo en la primera hora de vida. La fotografía pertenece a una serie encargada por John Lind, del Instituto Carolino de Estocolmo, para un estudio de recién nacidos.

Interpretación de la trompeta barroca

Los estudios históricos y físicos de ese olvidado instrumento permitirán que los trompetistas contemporáneos lo toquen como lo hicieran antaño los músicos de los siglos XVII y XVIII

Don Smithers, Klaus Wogram y John Bowsher

El actual renacimiento de la música barroca pone especial atención en los métodos de canto e interpretación instrumental que satisfacen los criterios de una intepretación correcta desde la perspectiva histórica. Muchos conjuntos instrumentales europeos y norteamericanos se dedican a la práctica de la música de los siglos XVII y XVIII aplicando técnicas auténticamente barrocas, con instrumentos originales, o modernos copia de aquellos que han sobrevivido. Las dificultades que han debido superar son realmente formidables. Tanta o más investigación científica e histórica exige la restauración de una obra musical que la de un cuadro o un edificio antiguo.

Hoy han cobrado nueva vida la mayoría de los instrumentos barrocos; de los supervivientes originales se han extraído prácticamente todos los secretos. Sin embargo, en unos pocos, esa revitalización hasta alcanzar un nivel de interpretación agradable al oído y a la vez correcto desde el punto de vista histórico ha sido lenta. De ellos, la trompeta barroca es la que peor se conoce, al parecer por dos razones fundamentales, que podrían ser complementarias. La revitalización de un instrumento antiguo depende de la continuidad de su diseño y de su construcción, así como de la persistencia de sus técnicas de interpretación en la memoria de los músicos. Si existió un período significativo durante el cual dejó de fabricarse el instrumento, con la correspondiente pérdida de la maestría necesaria para tocarlo, con poco fundamento cabrá esperar que su revitalización resulte rápida o fácil. La trompeta barroca no se fabrica hoy como antaño, en los siglos XVII y XVIII, y sus métodos de interpretación han sufrido graves componendas.

A diferencia de la trompeta moderna, el instrumento del barroco no disponía de elementos mecánicos (pistones, llaves, agujeros y similares), al objeto de obtener una escala de tonos completa, desde su registro más grave hasta el más agudo. Se trataba simplemente de un tubo metálico largo, arrollado en espiral o doblado, muy parecido a la corneta moderna. Al igual que ésta, contenía una columna de aire de longitud total fija. Por tanto, el instrumento se encontraba hasta cierto punto limitado a aquellas notas que pueden producirse a partir de una columna de aire fija, los denominados armónicos o sobretonos. Por ese motivo se habla a veces de las trompetas y las trompas barrocas como de instrumentos naturales.

El trompetista barroco, limitado como estaba a los tonos naturales –los armónicos– debía disponer de un instrumento de longitud y diseño apropiados y desarrollar además una técnica especial de interpretación que lograra extraer suficientes armónicos para generar una escala musical. La longitud de ese instrumento sería de unos 240 centímetros (un tubo más corto daría menos armónicos útiles). El tubo de esa medida tiene como nota más grave, o fundamental, el do^{1} (dos octavas por debajo del do central del piano). Así, tal trompeta era capaz de producir una serie discreta de armónicos (denominados parciales) por encima de la nota fundamental, que también recibe el nombre de primer armónico [*véase la figura 4*].

Puede alterarse la fundamental modificando la longitud del tubo, con lo que se obtiene una serie más aguda o más grave de armónicos superiores, según sea la frecuencia de la nueva fundamental. El trombón constituye un ejemplo de instrumento de metal con fundamental variable, pues dispone de un sistema de corredera que modifica la longitud de la columna de aire. El trombonista elige notas de distintas series armónicas y, para formar una escala completa, rellena los huecos de una serie particular. En la trompeta y en la trompa modernas los pistones desempeñan una función muy parecida a la de la corredera del trombón.

Uno de los problemas que presenta la trompeta barroca es que varias notas de su serie armónica no cumplen los criterios de la afinación occidental. Le complica ello las cosas al trompetista barroco y constituye el principal obstáculo para recuperar el arte de la interpretación de ese instrumento. Los armónicos naturales, entre el sexto y el décimosexto parcial, que son números primos (sólo divisibles por sí mismo, y por 1) están claramente desafinados respecto de cualquier norma de temperamento musical occidental. El séptimo armónico (y por tanto su octava) es demasiado grave; el parcial 11 ni es fa^{4} ni fa sostenido^{4}, el parcial 13 ni es la^{4} ni sol sostenido^{4}.

Es más, algunas notas, como re^{3}, fa^{3}, la^{3}, si natural^{3}, do sostenido^{4}, y mi bemol^{4} que aparecen con frecuencia en las partituras escritas para el instrumento por compositores de los siglos XVII y XVIII, no aparecen en absoluto en la serie armónica que tiene como fundamental al do^{1}. Sin embargo, se supone que sí podía darlas el maestro trompetista barroco, y ello sin contar más que con el instrumento habitual de altura fija que, en circunstancias normales, sólo podía producir la serie de armónicos que se ha visto antes. Hasta que se demostró, tanto en conciertos como en grabaciones, que en verdad



1. JOHANN GOTTFRIED REICHE, trompetista barroco, con su *tromba da caccia* espiral. Músico municipal de número, fue primer trompeta de la orquesta de J. S. Bach, en Leipzig, desde 1723 hasta su muerte en 1734. La partitura que sostiene en la mano izquierda es una fanfarria, breve pero muy difícil, que ha grabado uno de los autores (Smithers) utilizando una réplica del instrumento y aplicando las técnicas de ejecución propias del barroco. La CBS emplea esa grabación como sintonía de uno de sus noticiarios. El retrato, obra de Elias Gottlob Haussman, se encuentra en el antiguo ayuntamiento de Leipzig.

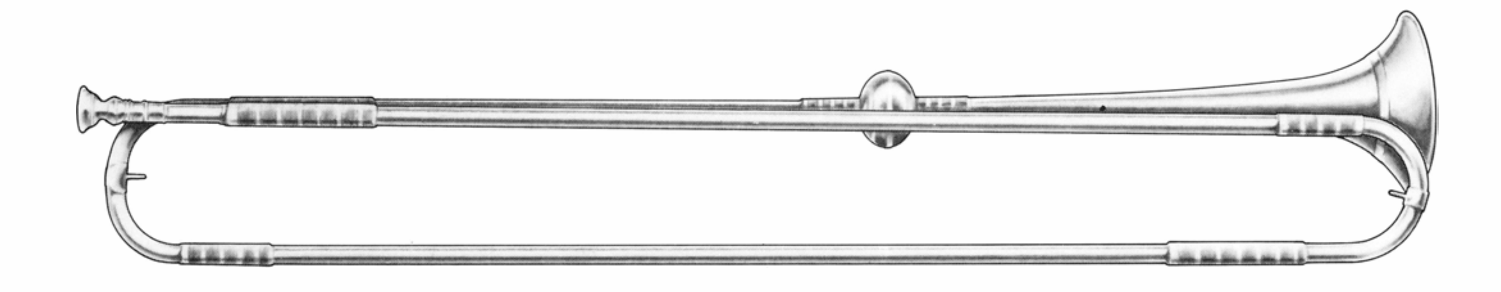
podían ejecutarse esas notas con la afinación correcta y la suficiente resonancia en una auténtica trompeta barroca, muchos creyeron que se trataba de errores o que estaban pensadas para un instrumento de altura variable, como la trompeta con corredera. Demostraciones recientes efectuadas con instrumentos históricos de autenticidad probada atestiguan que los compositores barrocos sabían lo que hacían.

El último autor que, a lo que parece, tuvo una experiencia de primera mano con los temidos problemas de la interpretación de la trompeta barroca fue el músico alemán Johann Ernst Altenburg. Según dejó escrito, los trompetistas debían recibir formación teórica y práctica, "en parte debido a las notas que faltan y a las desafinadas, y en parte debido a que ese instrumento sigue constituyendo un misterio, mucho más que otros". Recientes estudios científicos –los primeros de este tipo– han intentado arrojar luz sobre el misterio y recabar datos que mejoren la comprensión de las dificultades técnicas que entraña la interpretación, a la trompeta barroca de altura fija, de innumerables partituras de Purcell, Bach, Händel y otros muchos compositores barrocos. Destacan esos estudios el nexo inseparable que une la ciencia con la práctica.

Los estudios anteriores solían concentrarse en las características acústicas del instrumento. Se prestaba muy poca atención a la integración del instrumento, su boquilla y el músico. Para entender los misterios de la interpretación de la trompeta barroca se requiere un examen de la interacción dinámica entre los tres componentes, más todavía cuando sabemos hoy que el intérprete ejerce un gran efecto sobre la producción de sonido de un instrumento, modificando notablemente su eficacia acústica. Los labios del músico, los dientes, la lengua, la cavidad bucal y la garganta influyen también directamente sobre el sonido, sean cuales fueren las características físicas del propio instrumento.

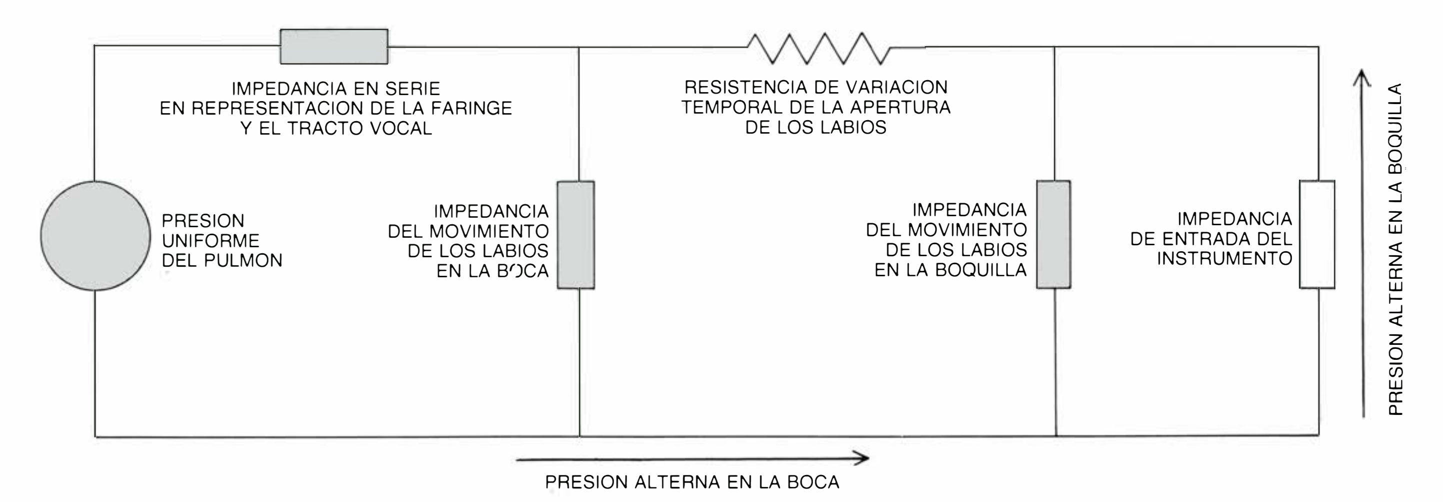
De importancia comparable es lo que sucede en las zonas situadas antes (detrás) y después (delante) de los labios sometidos a vibración: los ambientes previbratorios y postvibratorios no se excluyen mutuamente, sino que interactúan. En otras palabras, la generación del sonido de la trompeta no depende exclusivamente de la región de aire vibrante que, desde los labios del músico, se extiende hasta el extremo del instrumento, en forma de pabellón. La vibración de los labios imprime al aire regímenes de oscilación dentro de la boquilla y del instrumento, pero también en la cavidad bucal del músico. Se está estudiando cómo interactúan esas resonancias. Los primeros datos señalan que la interacción, hasta ahora inexplicable, constituye un factor importante en la generación y control de ciertas notas, especialmente en los registros medios y agudos del instrumento.

Por ejemplo, cualquier intérprete de un instrumento de metal sabe que, al dar notas prolongadas, una ligera manipulación de la lengua o de los músculos de la garganta, o de ambos a la vez, permite la ejecución de trinos sin recurrir a sistemas mecánicos (pistones o llaves). Aun sin entender la base teórica subyacente, el músico con experiencia modifica así las resonancias postvibratorias, que a su vez controlan lo que se oye del intérprete y de su instrumento. Es incorrecto, por tanto, denominarlos trinos labiales, como se hace habitualmente: las fluctuaciones de las resonancias anteriores a los labios del músico (en las regiones previbratorias) se bastan para modificar la dinámica de la columna de aire postvibratoria (que ocupa la boquilla y el instrumento). Se logran efectos similares desplazando el propio aparato vibratorio, bien sea golpeando la boquilla contra los labios o por ciertos movimientos de la mandíbula inferior. Sin embargo, tales métodos no suelen ser satisfactorios; no alcanzan el tipo de control que ejerce la fluctuación sutil de las resonancias previbratorias en la boca del músico y en el tubo.





2. DOS TIPOS de trompeta barroca. Arriba, una trompeta natural, larga y doblada dos veces (sin válvulas ni pistones) afinada a un re agudo. Este instrumento pertenece a un juego de tres trompetas construidas por Johann Leonhard Ehe en Nürnberg, en 1746. El otro instrumento es una *tromba da caccia* en espiral afinada a un re grave. La construyó Johann Wilhelm Haasen Nürnberg, en 1688. La empresa Haas de constructores de trompetas, trompas y trombones tuvo fama de producir las mejores trompetas naturales de la era barroca. La *tromba da caccia* que aparece en el retrato de Reiche que muestra la figura 1 bien pudo ser una trompeta Haas



3. DIAGRAMA ELECTRICO que representa la integración de intérprete, boquilla e instrumento de metal. Los pulmones del músico, en serie con la faringe y el segmento vocal, representado aquí por una impedancia en serie, generan una presión uniforme. Esa impedancia contiene todos los ajustes que quedan bajo control del músico, como la apertura de las mandíbulas y la posición de la lengua. Los labios del músico se acercan y alejan de la boquilla, provocando flujos de volumen en respuesta a la presión en esas dos posiciones, representadas como impedancias del movimiento de los labios. La resistencia de variación temporal de la apertura de los labios se hace infinita cuando los labios están cerrados; adquiere valores pequeños cuando los labios se encuentran muy abiertos, en la interpretación de notas graves. Esa gran variación cíclica se traduce en las características formas de onda de presión que se observan en la boquilla. Se muestran aquí como presión desarrollada en la boquilla a través de la impedancia de entrada del instrumento y presión del aire medida en la boca del músico.

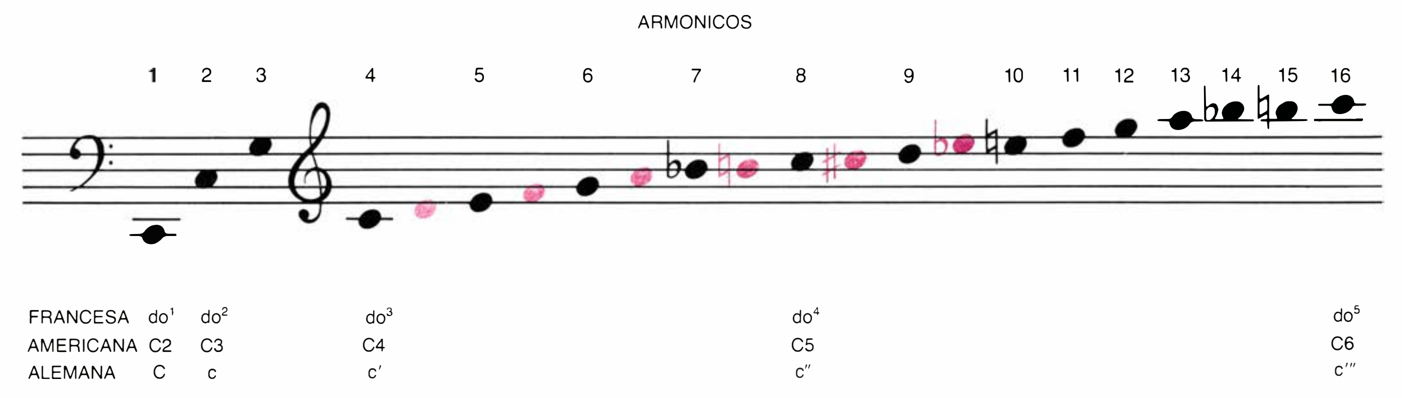
A diferencia del resto de instrumentos musicales, los de metal no poseen oscilador propio. El sonido lo producen sólo las vibraciones de los labios del músico, que a su vez modulan el aire contenido en todo el sistema. De ahí la influencia característica del músico en diversos parámetros acústicos, como la sonoridad y la altura de esos instrumentos.

La integración de músico, boquilla e instrumento puede representarse, en un diagrama eléctrico, mediante un modelo análogo simplificado de esos tres componentes característicos [*véase la figura 3*]. Sostiene la teoría que para que oscilen los labios de un trompetista las impedancias postvibratorias y previbratorias totales deben tener igual valor, pero fase opuesta. En un instrumento resulta fácil satisfacer esa condición, y así el músico puede controlar el tipo exacto de oscilación valiéndose de pequeños ajustes de las impedancias previbratorias. Sin embargo, cuando el músico hace vibrar los labios, solos o únicamente con la boquilla, las impedancias postvibratorias corresponden a la radiación libre o a las de la boquilla. En esa situación, el músico se ve obligado a efectuar mayores ajustes de las impedancias para asegurar que se produzca la vibración. Una sencilla estimación de las magnitudes que intervienen destaca la importancia de las impedancias cuyo control corresponde al intérprete y la poca importancia de la impedancia del instrumento.

El parámetro peor comprendido en la interpretación de los instrumentos de metal, en general, y de la trompeta barroca en particular, es la dinámica de los labios del músico. Dado que les corresponde a éstos la contribución principal (e inicial) a la generación del sonido de cualquier instrumento de metal, resulta paradójico y lamentable ese desconocimiento de la dinámica de sus vibraciones. Lo que en este análisis se aborda es la relación no lineal que guardan el flujo de aire que atraviesa un orificio y la presión que se registra a un lado y otro del mismo.

El orificio formado por el intersticio de los labios se reduce a medida que el músico da notas más agudas. Con el incremento de la frecuencia se produce un aumento de la resistencia, rápido y desproporcionado, mientras que, simultáneamente, se reduce la masa vibrante, lo que a su vez se traduce en una reducción de la superficie de la fuente vibratoria y un acortamiento de la amplitud.

Consecuencia de esos fenómenos es la diferencia inherente entre las notas tocadas con una boquilla grande y las que se dan con una boquilla pequeña.



4. SERIE DE ARMONICOS basada en la fundamental do[1] (dos octavas por debajo del do central del piano, do[3]). Muestra la gama de notas accesible al intérprete de una trompeta barroca de un tubo de 240 centímetros. El instrumento podía producir una serie limitada de parciales (*negro*), o tonos armónicos, por encima de la fundamental. Ciertas notas (*color*) aparecen con frecuencia en las partituras escritas para trompeta barroca en los siglos XVII y XVIII, pero no forman parte de la serie armónica del do[1]. A lo que parece, se exigía al trompetista la ejecución de esas notas pese a que su instrumento era de altura fija.

Refrenda esta conclusión la experiencia de tocar la trompeta barroca en su versión original. Con una boquilla grande, de bordes planos, se alcanzan notas del registro agudo de clarín (a partir del armónico 12 en adelante) donde la superficie de vibración de los labios difiere de la correspondiente a las notas graves del registro *basso*. Cabe observar que la longitud de la columna de aire de la trompeta barroca era aproximadamente igual a la del trombón moderno y su boquilla tenía dimensiones similares y, sin embargo, el registro clarín del instrumento es, por lo menos, una octava más agudo que las notas más agudas que deben ejecutar hoy la mayoría de trombonistas.

En poco se parecen las técnicas habituales de los trompetistas modernos con las artes de que debían valerse los del barroco para emitir esos agudos. Las boquillas modernas son mucho menores que las construidas en el siglo XVII y principios del XVIII, de ahí que sólo se obtengan notas agudas apretando o aplastando los labios al forzarse el aire a través del tejido subcutáneo, apretando con gran tensión. Provoca ello una rápida fatiga de los labios, e incluso la lesión de tejido labial, como atestiguará más de un músico.

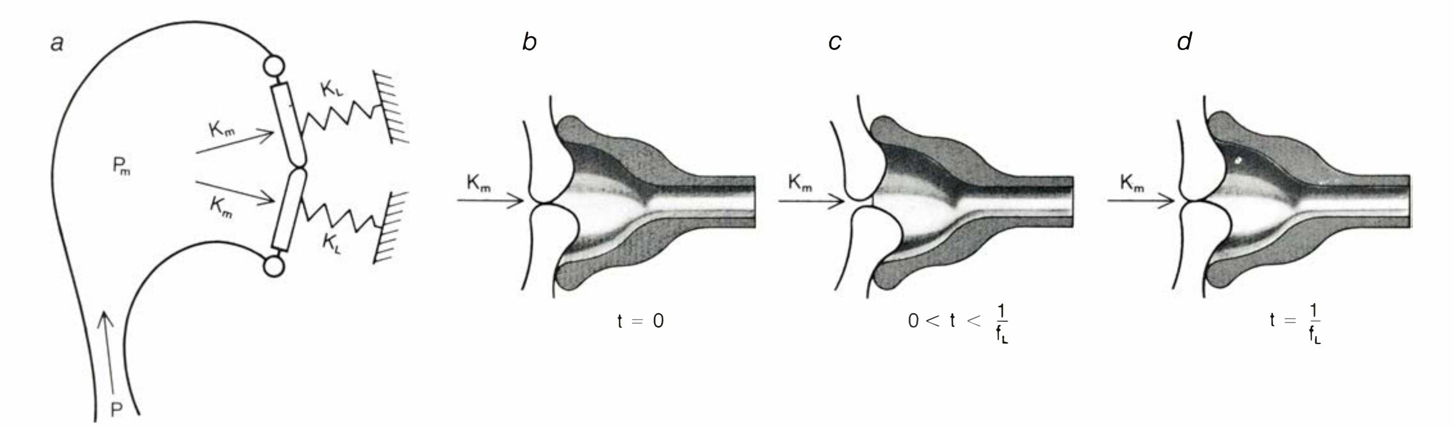
Los labios son un tejido vivo, y constituyen la única fuente de sonido en los instrumentos de metal; existe, por ello, gran correlación entre la interpretación de la trompeta barroca y el canto. Los labios del trompetista funcionan como lo hace la laringe de un cantante, de ahí la fuerte semejanza entre la forma de tocar un clarín y las técnicas utilizadas por una soprano coloratura. En efecto, muchos textos sobre la interpretación de instrumentos de metal afirman que el alumno debe tomar lecciones de canto. Aconsejaba un autor del siglo XVIII que el intérprete de clarín imaginara siempre estar cantando, y que procurara imitar una voz lo más bella posible.

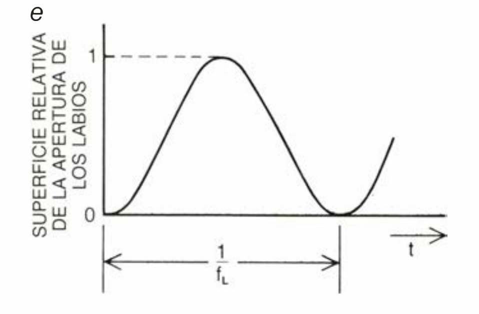
Muchos intérpretes de trompa, trombón o trompeta conocen la correlación que guardan el canto de las vocales primera y tercera (*a* e *i*) y la obtención de notas desde el registro más grave (*a*) hasta el más agudo (*i*) en los instrumentos de metal. La trompeta barroca parece especialmente sensible al menor cambio de las resonancias previbratorias, bien sea por la longitud del tubo del instrumento, de aproximadamente 240 centímetros, o por la combinación de varios factores acústicos, incluidos los inherentes a la boquilla.

Para que se produzcan notas en el registro *basso* inferior, el paso de aire debe estar abierto, la lengua bajada en la posición de la primera vocal y los labios lo más sueltos e hinchados posible. Las notas agudas, por el contrario, requieren que la lengua se arquee hacia arriba, apretando el paladar (en la posición de la tercera vocal), dejando el mínimo espacio al paso del aire y cerrando casi el paso desde la cavidad bucal. Además, los labios, apretados, deben arrugarse.

El examen estetoscópico de los instrumentistas de trompeta barroca tocando desde el registro más grave hasta el más agudo ha verificado el efecto de esas resonancias previbratorias. Se observó que las notas graves se tocaban con mayor intensidad en la región de la laringe y en los bronquios superiores, mientras que las frecuencias más agudas alcanzaban su máxima amplitud en las regiones superiores de la garganta y debajo de la mandíbula. Los sonidos captados en las mejillas y en los huesos cercanos a la nariz quizá los genere la vibración del aire en la cavidad bucal, o la conducción de sonido desde los labios a los dientes y hacia el maxilar superior.

El cambio de la posición de la lengua afecta también al timbre, o color, de las notas que emiten los instrumentos de metal. Se observa ese efecto en la grabación de espectros sonoros en varias actividades comparables [*véase la figura 7*]. Cuando se graba la nota do^3 a una distancia de un metro frente al intérprete, que hace sonar los labios sin boquilla o sin instrumento, el efecto es virtualmente el mismo que cuando se emiten notas agudas y graves con una guimbarda. El número de sobretonos aumenta cuando se arquea la lengua en la posición de la tercera vocal, aguda, (lo que reduce el volumen de resonancia de la boca) y disminuye cuando se





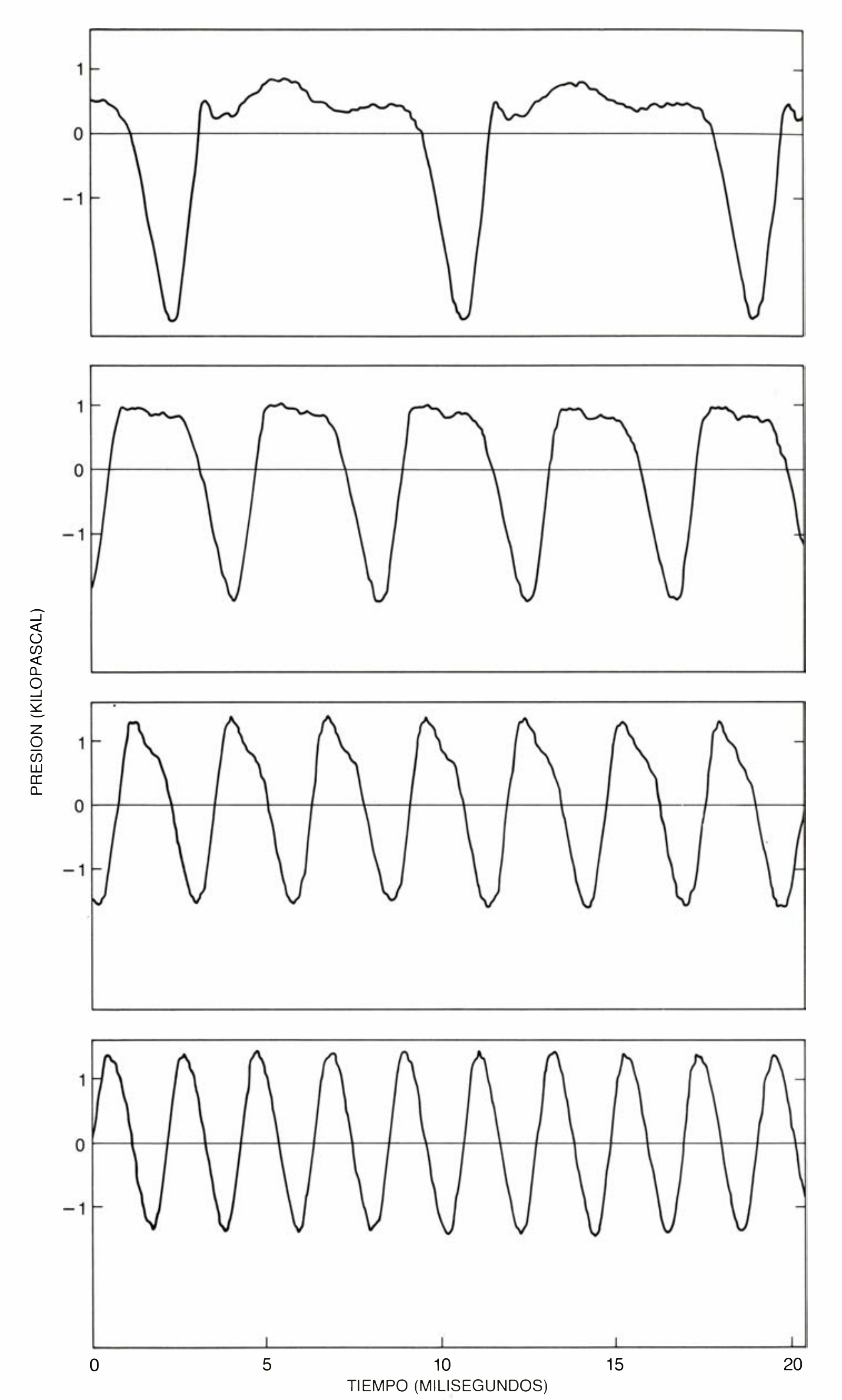
5. ACTUACION DE LOS LABIOS en un instrumento de metal, representada en un diagrama esquemático donde *P* es la presión de aire en los pulmones, P_m la presión de aire en la boca del músico, K_L la fuerza muscular en los labios, K_m la fuerza del aire comprimido contra los labios, f_L la frecuencia de oscilación de los labios y *t* el tiempo. Los labios representan un sistema oscilante formado por una masa *m* y un muelle (*a*). Se comportan como dos aletas que tienden a mantenerse cerradas por la fuerza elástica que actúa sobre ellos (*b*). Al aumentar la presión interna en la boca, cuando la fuerza K_m supera a la fuerza K_L los labios se abren (*c*). Se escapa entonces cierta cantidad de aire en forma de impulso positivo de presión. La reducción de presión resultante hace que los labios se cierren (*d*); aumenta de nuevo la presión, que fuerza la reapertura de los labios. Se constituye así un régimen particular que muestra una variación sinusoidal de la superficie de apertura de los labios con respecto al tiempo (*e*). Si la masa de los labios, la tensión de éstos y la presión de aire se mantienen constantes, el tiempo de relajación disminuirá al reducirse el volumen de la cavidad bucal. El músico puede variar el volumen alterando la posición de la lengua. Facilita la producción de notas graves el descenso de la lengua y, de notas agudas, el arqueo de la lengua hacia arriba y adelante.

baja la lengua a la posición de la primera vocal (aumetándose el volumen de resonancia de la boca). Simultáneamente crecen los componentes de ruido, según prueba el incremento de 20 decibelios del sonido de frecuencia superior a los seis kilohertz (6000 ciclos por segundo). Lo que significa que la frecuencia de resonancia de la cavidad bucal aumenta al levantar la lengua, y que se registra más turbulencia en la circulación del aire y una disminución del desplazamiento del labio.

Se obtienen espectros comparables cuando, en lugar de limitarse a vibrar los labios, el músico sopla en una trompeta natural en idénticas condiciones [*véase la figura 8*]. Si el volumen de la boca es grande (con la lengua bajada) el sonido aparece lleno y brillante, con gran riqueza de sobretonos y muy poco ruido, producto de la intensa vibración y el gran desplazamiento de los labios. Cuando disminuye el volumen de la boca, al elevar la lengua, el sonido resulta más áspero, de desagradable color nasal. Para obtener tonos estables en el registro muy agudo de clarín debe reducirse en lo posible el volumen de la boca, aunque no siempre se obtenga un timbre satisfactorio.

Después del instrumento propiamente dicho, la boquilla constituye el elemento de mayor importancia en la interpretación de los metales. Es parte integrante del instrumento, tan importante como éste en la generación de máximos de impedancia específica (la relación entre presión y velocidad de volumen en una determinada superficie) y en la obtención de una curva de afinación favorable (la curva que refleja los valores de las frecuencias que se tocan), especialmente para las notas superiores al sexto armónico. Por tanto, boquilla e instrumento deben emparejarse. Es más, como hemos señalado, el intérprete también forma parte del sistema total, con lo que la boquilla debe emparejarse a la vez al músico y al instrumento. Constituye la interfase entre ambos.

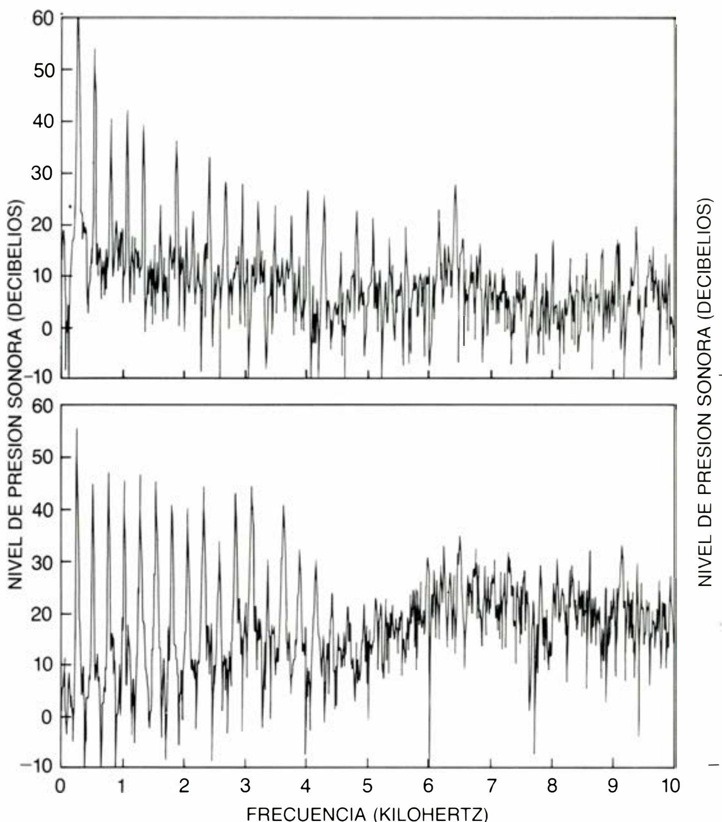
De hecho, y por esas razones, la boquilla es uno de los elementos más personales del músico. Tal relación parece más acusada en el caso de la trompeta barroca que en el de otros instrumentos de metal, y ello debido a la gama de notas que debe producir. El registro *basso* más grave del instrumento corresponde al del trombón. El registro clarín, más agudo, suele situarse por encima del registro normal de la trompeta moderna, que ya está una octava por encima del límite superior del trombón, y por lo menos una quinta por encima del de la trompa.

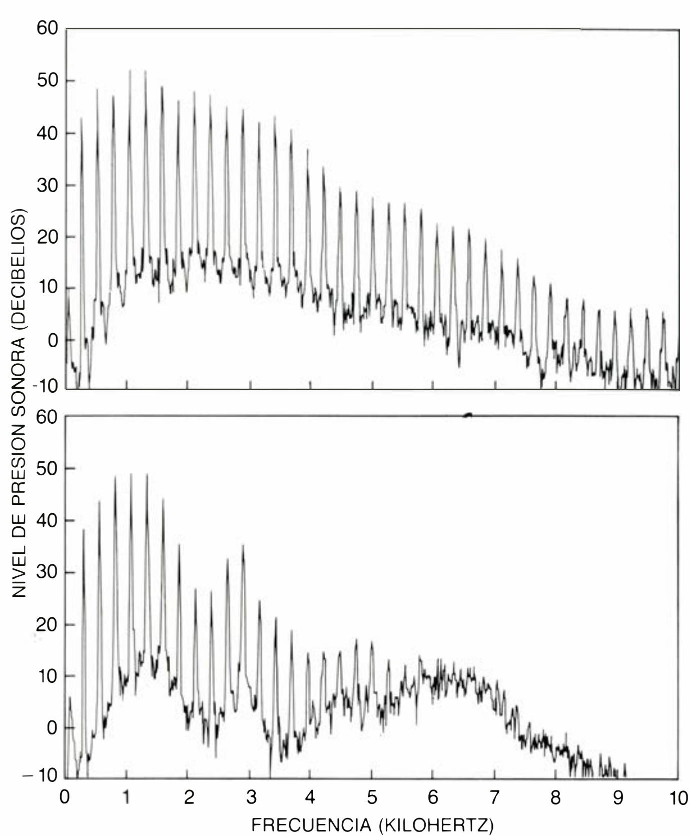


6. PRESION ALTERNA medida con un micrófono de sonda en la boquilla de un trombón al dar las notas si bemol^1, si bemol^2, fa^3 y si bemol^3; se muestran sus respectivas mediciones de arriba abajo. La línea del cero en la escala de presiones corresponde a una presión aproximadamente equivalente a la atmosférica.

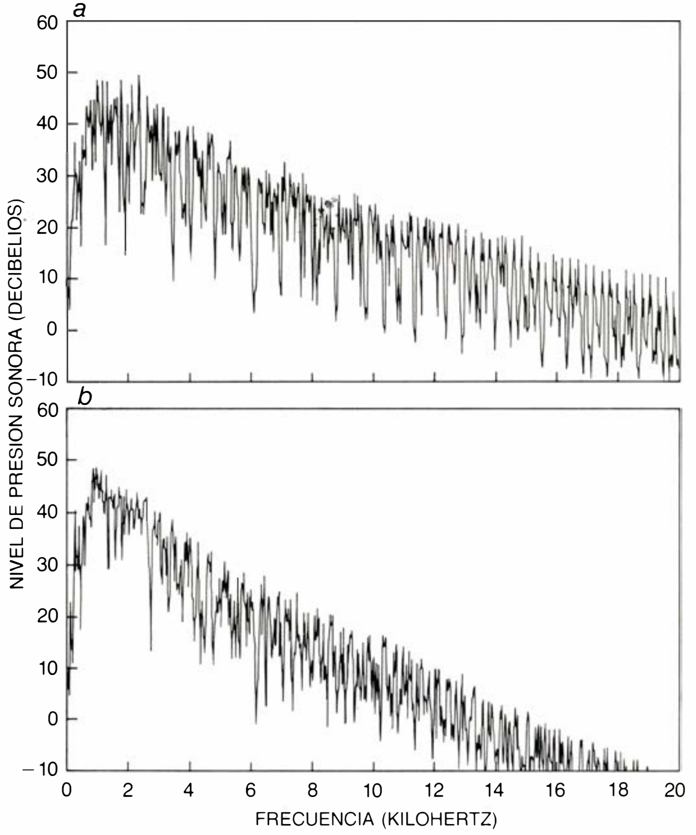
El gran número de boquillas de trompeta barroca que han llegado a nuestros días sugiere que intérpretes y constructores daban gran importancia a su diseño y fabricación. No hay dos iguales. No cabe duda de que se hacían según el gusto y la estructura facial de cada intérprete. Idéntica importancia debió tener el emparejamiento con el instrumento correspondiente.

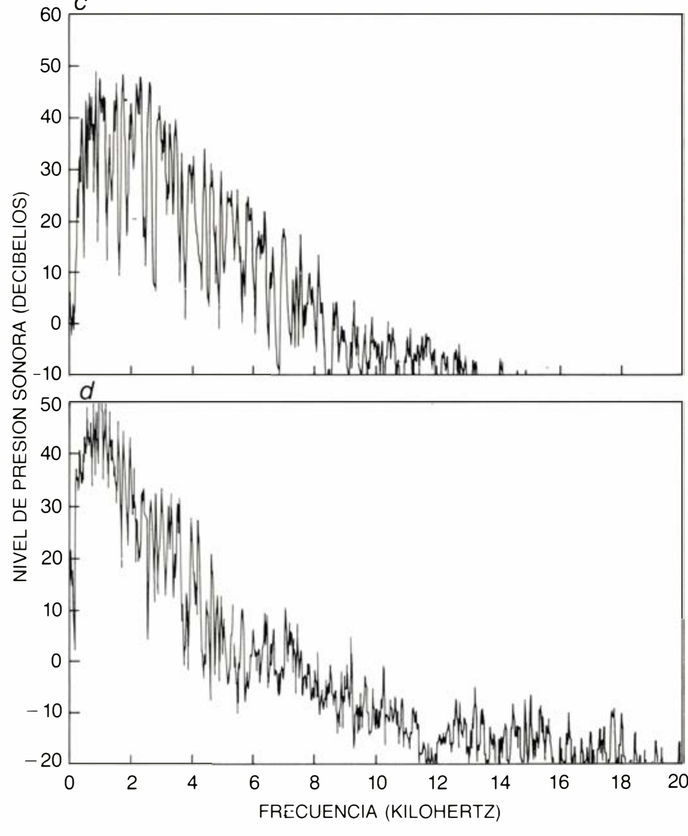
Los músicos actuales saben adap-





7. ESPECTROS SONOROS de dos formas de ejecutar el do³. En los registros de la izquierda, el intérprete de trompeta barroca hace sonar exclusivamente los labios, sin boquilla ni instrumento. A la derecha toca una trompeta natural con un tubo doblado dos veces. En ambos casos el gráfico superior corresponde a un gran volumen interior de la boca (situada la lengua hacia atrás y abajo), y el inferior a un volumen de boca mínimo (lengua hacia delante y arqueada).





8. A CUATRO INSTRUMENTOS corresponden estos espectros sonoros, que muestran los valores medios en un intervalo de dos octavas. Los instrumentos son una trompeta barroca en do con tubo doblado dos veces (*a*), una trompeta barroca en espiral, la *tromba da caccia*, o clarín, en re (*b*), una trompeta piccolo moderna en si bemol (*c*) y una corneta en la (tocada con taladros) (*d*). Las diferencias entre las trompetas barrocas y la trompeta piccolo, que a menudo las sustituye en las interpretaciones modernas, demuestran que la música para trompeta barroca no alcanza el brillo y la claridad requeridos cuando se ejecutan en piccolo, cuya longitud es cuatro veces menor que la del instrumento en que pensaba el compositor al escribir la partitura. (Dibujos de George Retseck.)

tarse a boquillas de tipos diversos, y suelen emplear la misma durante largo tiempo. El proceso conlleva un cierto riesgo para el delicado aparato neuromuscular del generador de sonido (los labios y los músculos del cuello y de la cara). Si le cabe elección, el instrumentista siempre preferirá la boquilla que le resulte más confortable y dé la mejor respuesta acústica.

Pese al carácter irrepetible de cada boquilla de trompeta barroca, presentan en conjunto una serie de características comunes. Casi todas son mayores que las de mayor tamaño de las trompetas modernas. Según indican diversos procedimientos de datación, las boquillas del siglo XVII son mayores que las construidas en el XVIII. Más importante todavía, el diseño de las boquillas de trompeta barroca difiere básicamente del de las boquillas de la trompeta moderna. Facilita una mayor resonancia de los armónicos de frecuencias más graves, confiere mayor definición a los parciales más agudos y permite un mejor control de los tonos no armónicos.

Además, el diseño particular de algunas boquillas de trompeta barroca aumenta la excitación de las frecuencias más agudas. Lo cual, junto con la presencia, más profunda y resonante, de los armónicos más graves generados en la columna de aire de 240 centímetros da a cada tono un espectro sonoro más rico y amplio que en las trompetas modernas, de menor longitud. La moderna trompeta piccolo a pistones, la más utilizada en las interpretaciones contemporáneas de música barroca, produce notas algo estridentes, ya que sólo mide unos 60 centímetros.

El último término de la ecuación del toque de la trompeta barroca, no menos importante que los demás, es el propio instrumento. Nos sale aquí al paso una peculiaridad: pese a la perfección de los componentes modernos, construidos a máquina, frente a las irregularidades de la construcción artesanal del siglo XVIII, cuesta menos tocar los instrumentos antiguos, que además están mejor afinados que sus réplicas modernas. Recientes investigaciones acústicas han demostrado que la afinación de un instrumento (los valores de las frecuencias de resonancia) está determinada casi exclusivamente por la forma de las paredes interiores de la columna de aire, mientras que la respuesta de un instrumento depende en su práctica totalidad del denominado factor Q de la resonancia, esto es, la suavidad y el grado de continuidad de la superficie interior del tubo metálico (y en cierta medida del material que se haya empleado en la construcción del tubo).

Las trompetas barrocas, tanto las originales como las modernas, se han construido en su mayoría empalmando tubos cilíndricos (bien sea arrollados o doblados), con un final cónico en el pabellón. En las réplicas modernas los métodos de fabricación logran una superficie interior regular y suave, creándose resonancias con un alto factor Q. Dado que no existe una relación particular entre las frecuencias de resonancia y los componentes armónicos del sonido, los instrumentos modernos presentan menos variación de altura que los originales de los siglos XVII y XVIII, y cuesta tocarlos con la debida afinación.

Los componentes de los antiguos instrumentos de metal se fabricaban a partir de planchas de latón, igual que se hace hoy; las planchas martilladas a mano carecen, sin embargo, de la consistencia de las modernas planchas laminadas a máquina. Es más, cuando los artesanos fabricaban tubos moldeándolos y martillándolos sobre varillas de acero, introducían muchas irregularidades en la superficie interior. Esas imperfecciones, pequeñas pero significativas, se traducían en irregularidades del diámetro del tubo, juntas imperfectas y falta de simetría en los puntos de curvatura del tubo. Todas esas variaciones reducían el factor Q de las resonancias de la trompeta barroca, y por tanto aplanaban la curva de resonancia.

El resultado final de todo ello es que, en una trompeta barroca antigua, el músico puede "inflexionar", esto es, variar los armónicos naturales; así, muchos instrumentos antiguos permiten una interpretación que respete la afinación sin provocar alteraciones inaceptables del timbre o la respuesta. Para alcanzar similares prestaciones casi todos los fabricantes de trompetas barrocas modernas han optado por el recurso de practicar agujeros en el instrumento. Con los dedos se corrigen las alturas, respetando el músico la afinación. Ello no sólo corrompe los principios históricos, sino que establece un pacto derrotista entre los parámetros acústicos específicos y las técnicas de interpretación que les corresponden. Sólo aplicando principios históricos a los tres elementos de la ecuación –el músico, la boquilla y el instrumento– resucitará el arte, olvidado, de tocar la trompeta barroca.

Juegos de ordenador

Un programa que gira hipercubos induce demencia tetradimensional

A. K. Dewdney

"¡Mi marido se ha esfumado, y creo que tiene usted algo que ver en ello!" Era Cheyl quien me hablaba por teléfono; saltaba al oído que estaba furiosa. Al parecer, su marido, Magi, que es mi amanuense informático en la Universidad de Ontario Occidental, había desaparecido mientras observaba el funcionamiento de un programa que yo le había sugerido que escribiera. El programa hace girar un hipercubo –objeto tetradimensional análogo al cubo– y lo proyecta sobre una pantalla. Cheryl, muy excitada, prosiguió: "En el monitor hay una serie de líneas muy extrañas, y sus ropas están hechas un revoltijo junto a la silla. Tiene que haber estado usando esas gafas tan raras de cartón, esas que tienen cristales de colores. Y fíjese, ¡los calcetines siguen aún en los zapatos!"

He aquí, dije para mis adentros, un caso claro de demencia tetradimensional. Las víctimas de este mal quedan convencidos de que han logrado salir del espacio ordinario, y de que han penetrado en una dimensión superior, invisible para los demás. En algunos casos, la ilusión de haber desaparecido puede ser tan intensa que otros llegan a participar de ella: la víctima puede entrar en una habitación llena de gente y parecerle invisible a todos. Por fortuna, el caso de Magi tuvo desenlace feliz, pero lo guardaré para el final. En el ínterin, someteré el programa del hipercubo a un público más amplio, pero no sin hacer una llamada a la responsabilidad: los lectores susceptibles de ser presa de la demencia de Magi deben abstenerse de escribir el programa y, más aún, de verlo funcionar; insisto formalmente en ello. Entre las potenciales víctimas se encuentran todas aquellas personas con historial de obsesiones sobre espacios de dimensión mayor que tres, así como quienes se hayan sentido tentados –aunque sea ocasionalmente– por la expectativa de hallar realidades desconocidas.

La cuarta dimensión ha sido vehículo para la especulación física y metafísica desde el siglo XIX, cuando menos. La idea de una cuarta dimensión, provista de sentido físico, culminó con las teorías de la relatividad especial y general de Einstein; el espacio y el tiempo componen conjuntamente un continuo tetradimensional, en el cual se encuentran, intemporalmente congelados, todos los acontecimientos reales. Tal concepción del universo puede estar experimentando modificaciones dimensionales; las llamadas teorías Kaluza-Klein introducen siete o más nuevas dimensiones, bajo la forma de diminutas hiperburbujas ligadas a cada punto del espacio-tiempo [véase "Las dimensiones ocultas del espacio-tiempo", por Daniel Z. Freedman y Peter van Nieuwenhuizen; INVESTIGACIÓN Y CIENCIA, mayo de 1985].

Pero la cuarta dimensión que yo he llegado a conocer y amar es hija de la matemática. Los lectores, en una sala o habitación ordinaria, tienen suspendidos sobre sus cabezas un sistema de coordenadas tridimensionales. En cada rincón de la sala concurren tres paredes, y desde ese vértice parten tres rectas, cada una de las cuales es lugar de encuentro de dos paredes. Cada recta es perpendicular a las otras dos. ¿Podría el lector imaginar una cuarta recta que sea perpendicular a las otras tres? Seguramente no; pero ello es precisamente lo que exigen los matemáticos para establecer la construcción puramente conceptual que denominan de cuatro dimensiones. Los lectores tienen ahora la oportunidad de explorar este espacio de forma personal, y sin peligro para sus personas. Basta con que escriban el programa que llamo HIPERCUBO.

Los orígenes de HIPERCUBO se remontan a una película producida hace unos veinte años por A. Michael Noll, a la sazón en los Laboratorios Bell, que representan las sombras bidimensionales de objetos de cuatro dimensiones al moverse en el hiperespacio tetradimensional. Sin embargo, en su forma actual, el programa fue desarrollado por Thomas Banchoff y sus colegas, del laboratorio de gráficos informáticos de la Universidad de Brown; las fascinantes imágenes que genera me sirvieron de inspiración para este artículo [*véanse las figuras 1, 4 y 5*]. Banchoff, que es profesor de matemáticas, orienta la exploración visual de las superficies y espacios de dimensión superior a complementar sus investigaciones y escritos como geómetra. En 1978, Banchoff y Charles Strauss produjeron una película en color, de 9 minutos y medio de duración, generada por ordenador, titulada *The Hypercube: Projections and Slicing*, que se ha convertido desde entonces en clásico del *underground* matemático. Probablemente sea Banchoff también el más distinguido especialista en la vida y obra de Edwin A. Abbott, clérigo y maestro inglés que en 1884 escribió *Planilandia*, cuento sobre una imaginaria vida en dos dimensiones. (Existe edición castellana; véase "Bibliografía".)

Banchoff y sus colegas han concebido llamativas imágenes, que ilustran las propiedades de los objetos tetradimensionales. Así, por ejemplo, las imágenes de la figura plasman la rotación en el espacio de cuatro dimensiones de un hipercubo tetradimensional. Para comprender y apreciar las imágenes, fijémonos en las sombras que un cubo ordinario proyecta sobre un plano. La sombra puede a veces recordar un cuadrado dentro de un cuadrado. Coloreando con un matiz adecuado las caras del cubo, la sombra semejará un cuadrado con un agujero cuadrado en su interior [*figura 3*].

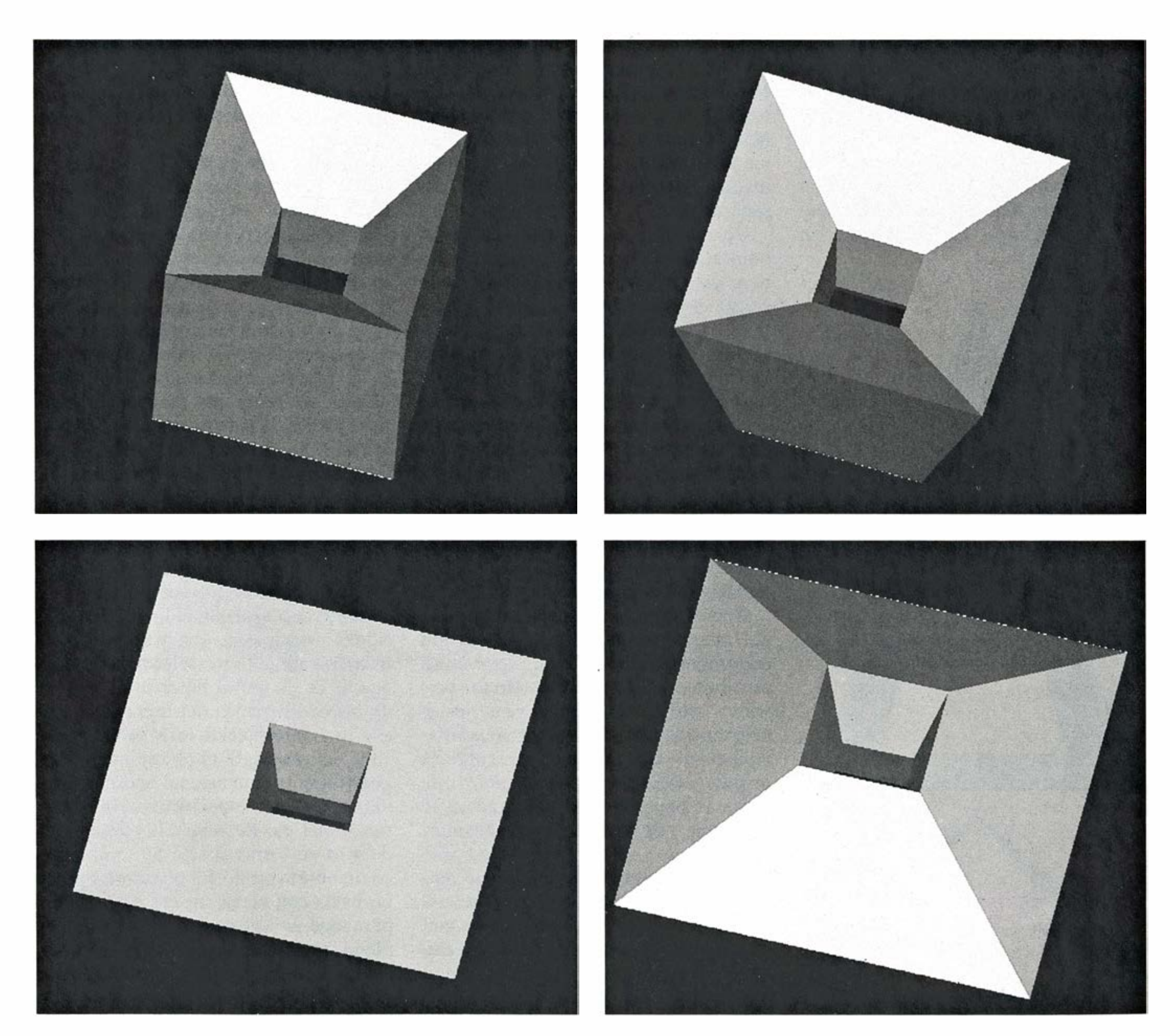
Análogamente, al iluminar un hipercubo desde un punto situado "por encima" del espacio ordinario, en la cuarta dimensión, la "sombra" tridimensional proyectada por el hipercubo puede asemejarse a un cubo con otro cubo en su interior. El cubo interior está rodeado por seis poliedros de seis caras, que podemos considerar cubos deformados. Los cuatro cubos adyacentes a los lados del cubo interior se acoplan entre sí y forman el sólido cuya superficie recuerda la de un toro compuesto por cajas, que vemos en las imágenes de Banchoff. Los otros dos cubos deformados, el cubo interior y el exterior, forman también un toro macizo, que no se muestra en la figura. Conforme va girando el hipercubo, el agujero cuadrado del toro visible parece moverse hacia el observador. Quienes preparen el programa HIPERCUBO verán cambios similares, si bien no con tanta continuidad y realismo.

Las figuras 4 y 5 pertenecen a una película que próximamente presentarán Banchoff y sus colegas Hüseyin Koçak, David Laidlaw y David Margolis: *The Hypersphere: Foliation and Projections*. La hiperesfera es un objeto mucho más complejo que el hipercubo, y no lo describiré aquí detalladamente. No obstante, pueden comenzar a comprenderse las imágenes por analogía con una esfera ordinaria. Supongamos que inicialmente la esfera descansa sobre un plano tangente a su polo sur, y que a su polo norte inicial se haya fijado un punto luminoso. Las sombras proyectadas sobre el plano por los paralelos formarán una serie de círculos concéntricos [*véase la figura 3*]. Manteniendo fija la luz y haciendo girar la esfera, las imágenes de los círculos pueden dejar de ser concéntricas; la imagen de cualquier circunferencia que pase por la fuente de luz será una línea recta.

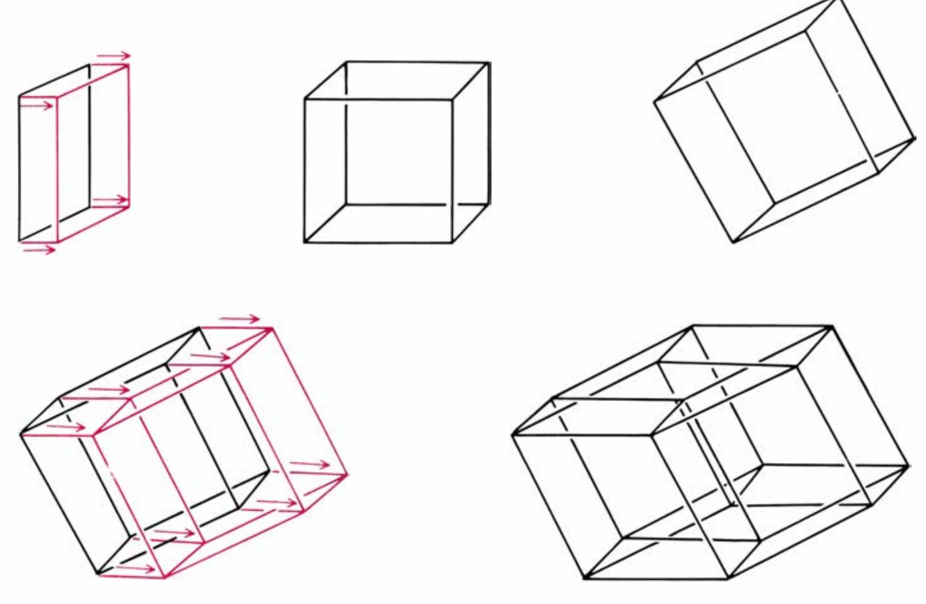
Análogamente, puede considerarse que la "sombra" tridimensional proyectada por una hiperesfera está formada por una serie de toros concéntricos [*véase la figura 4*]. En las imágenes de Banchoff, para hacer más visibles los toros se ha eliminado del toro exterior una cinta imaginaria supuestamente arrollada en torno a él. Al hacer girar la hiperesfera los toros parecen hincharse y engullirse unos a otros. Cualquier toro que pase a través de la fuente de luz se hace infinitamente grande [*véase la figura 5*].

Las analogías dimensionales son valioso instrumental para construir y comprender fenómenos de cuatro dimensiones. El hipercubo, por ejemplo, se deduce del cubo análogamente a como el cubo deriva del cuadrado. Para obtener el cubo a partir del cuadrado se alza el cuadrado en dirección perpendicular a su plano, hasta situarlo sobre él a una altura igual al lado [*véase la figura 2*]. El nuevo cubo tiene ocho vértices, el doble que el cuadrado de partida, y 12 aristas, cuatro debidas al cuadrado inicial, cuatro del cuadrado final sobrealzado sobre el inicial y otras cuatro resultantes de conectar los vértices del cuadrado inicial con sus homólogos del situado sobre él. El cubo tiene también seis caras cuadradas, una coincidente con el cuadrado inicial, otra con el cuadrado final y una erecta entre cada uno de los cuatro pares de lados que componen los cuadrados inicial y final.

Si por un momento admitiéramos

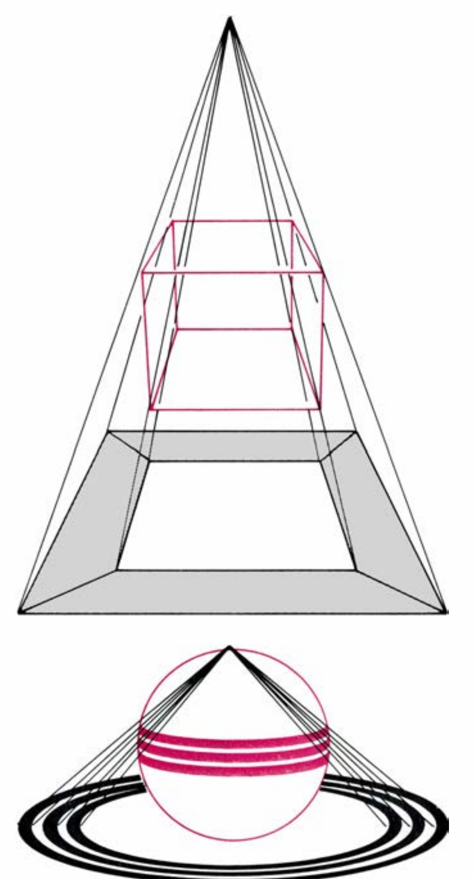


1. Proyecciones, sobre el espacio tridimensional, de un hipercubo tetradimensional que gira en torno a los ejes 2 y 4



2. Un cuadrado plano engendra un cubo; un cubo, un hipercubo

que estuviera disponible una dimensión más, podríamos repetir con el cubo la operación anterior: "alzaríamos" el cubo por encima del espacio ordinario, desplazándolo en la dirección de la nueva dimensión, hasta una distancia igual al lado del cubo [*véase la figura 2*]. Se obtiene así un hipercubo. ¿En qué dirección yace la nueva dimensión? No sabría explicarlo. Incluso una fotografía mía, en la que señalase hacia el interior de la cuarta dimensión, sería manifiestamente inútil. Pura y simplemente, mi brazo parecería haber desaparecido.



3. Proyecciones del cubo y la esfera

No obstante, poco cuesta contar el número de vértices, aristas, caras e hipercaras (cubos ordinarios) de que consta el hipercubo. El número de vértices es, sencillamente, el número de vértices del cubo inicial más el número de los del cubo final, 16 en total. Cada uno de los ocho vértices del cubo inicial estará unido por una arista con exactamente uno de los ocho vértices del cubo final; hay, además, 12 aristas en cada uno de los dos cubos. Así pues, el hipercubo tiene 8 + 12 + 12 = 32 aristas. Se puede demostrar también que el hipercubo tiene 24 caras planas ordinarias y ocho hipercaras.

Debo a David Laidlaw la explicación del programa HIPERCUBO. La versión del programa que describiré representa un hipercubo que muestra sólo sus vértices y aristas. Además, la vista que el programa genera no muestra necesariamente un cubo dentro de un cubo; la imagen dependerá, en cambio, de cómo se haya llevado a la práctica el programa y de cómo se haga funcionar. Cada vez que en el programa se hace girar el hipercubo, los vértices se mudan a nuevas posiciones y se genera una nueva vista del objeto, desazonadoramente desconcertante. Sin embargo, al cabo de una experimentación prolongada, las vistas comienzan a adquirir una extraña inteligibilidad, y uno se siente en el umbral de algo incitante, cuya amplitud infunde respetuoso temor.

Los 16 vértices del hipercubo manejado en el programa se numeran de 0 a 15, de acuerdo con un sencillo plan. Al escribir cada número en forma binaria y convertirlo en una matriz de cuatro dígitos binarios emerge un sistema de coordenadas en miniatura. Los dígitos binarios de 13, pongamos por caso, son 1 (es decir, una vez 8), 1 (una vez 4), 0 (ninguna vez 2) y 1 (una vez 1). El número binario puede denotarse entonces mediante la matriz (1, 1, 0, 1), que casi nos da un práctico sistema de coordenadas para la posición inicial del hipercubo. (No es una posición que se asemeje a un cubo dentro de otro.) Para convertir esta matriz binaria en un sistema de coordenadas útil se cambia cada 0 por −1, y se multiplica cada número de la matriz por otro lo bastante grande para engendrar una imagen de tamaño adecuado sobre la pantalla de presentación gráfica del ordenador. Por ejemplo, si el multiplicador fuese 10, las coordenadas del vértice número 13 serían (10, 10, −10, 10).

Parece como si las dimensiones se nos fueran infiltrando por doquier conforme vamos escribiendo HIPERCUBO. Una tabla o matriz bidimensional, llamada *vert*, conserva los vértices tal cual se definieron inicialmente. Dado que hay 16 vértices, de cuatro coordenadas cada uno, *vert* es una matriz de 16 por 4, que consta de 64 números; *vert*(i,j) es la j-ésima coordenada del i-ésimo vértice. El programa HIPERCUBO mantiene intacta la matriz *vert* que se define al principio del programa y cuyo contenido se transfiere después a una segunda matriz de 16 por 4, llamada *cubo*. La matriz *cubo* puede considerarse de trabajo; su contenido van modificándolo continuamente los giros efectuados por el programa.

HIPERCUBO está dividido en tres secciones principales, que vienen tras la inicialización de *vert*: selección del giro que se desea dar al hipercubo, cálculo de las coordenadas del hipercubo tras el giro y presentación en el monitor del cubo así girado. Si el objeto que ha de girar fuese tridimensional podría determinarse el giro especificando la orientación del eje de rotación y el ángulo de giro en torno al eje. Sin embargo, en un objeto giratorio tetradimensional no basta con elegir un eje de rotación para que se determine el giro de un plano: recordemos que, dado un plano cualquiera, existen dos direcciones no equivalentes perpendiculares a él. Por otra parte, incluso en el hiperespacio de cuatro dimensiones sigue siendo

cierto, como sucede en el espacio ordinario, que un giro sólo puede afectar a dos dimensiones cada vez. Al hacer girar un objeto tridimensional, dos de sus dimensiones van yendo de una a otra, mientras la tercera permanece fija. Análogamente, al girar un objeto de cuatro dimensiones, dos de ellas cambian de dirección en el espacio, mientras otras dos permanecen fijas.

Existen otros muchos modos de girar un objeto tetradimensional hasta situarlo en una nueva posición. Resulta, sin embargo, que puede alcanzarse cualquier posición aplicando una sucesión de giros que se limiten a inducir movimientos en el seno de los planos definidos por los ejes de coordenadas del espacio tetradimensional circundante. En un espacio tetradimensional hay cuatro ejes de coordenadas, numerados, pongamos por caso, de 1 a 4 y, por ello, seis modos de combinarlos de dos en dos. Por consiguiente, en el espacio tetradimensional determinado por los cuatro ejes de coordenadas hay seis "planos coordenados": el plano 1-2, que es el plano determinado por los ejes 1 y 2; el 1-3; el 1-4; el plano 2-3; el 2-4 y, finalmente, el plano 3-4.

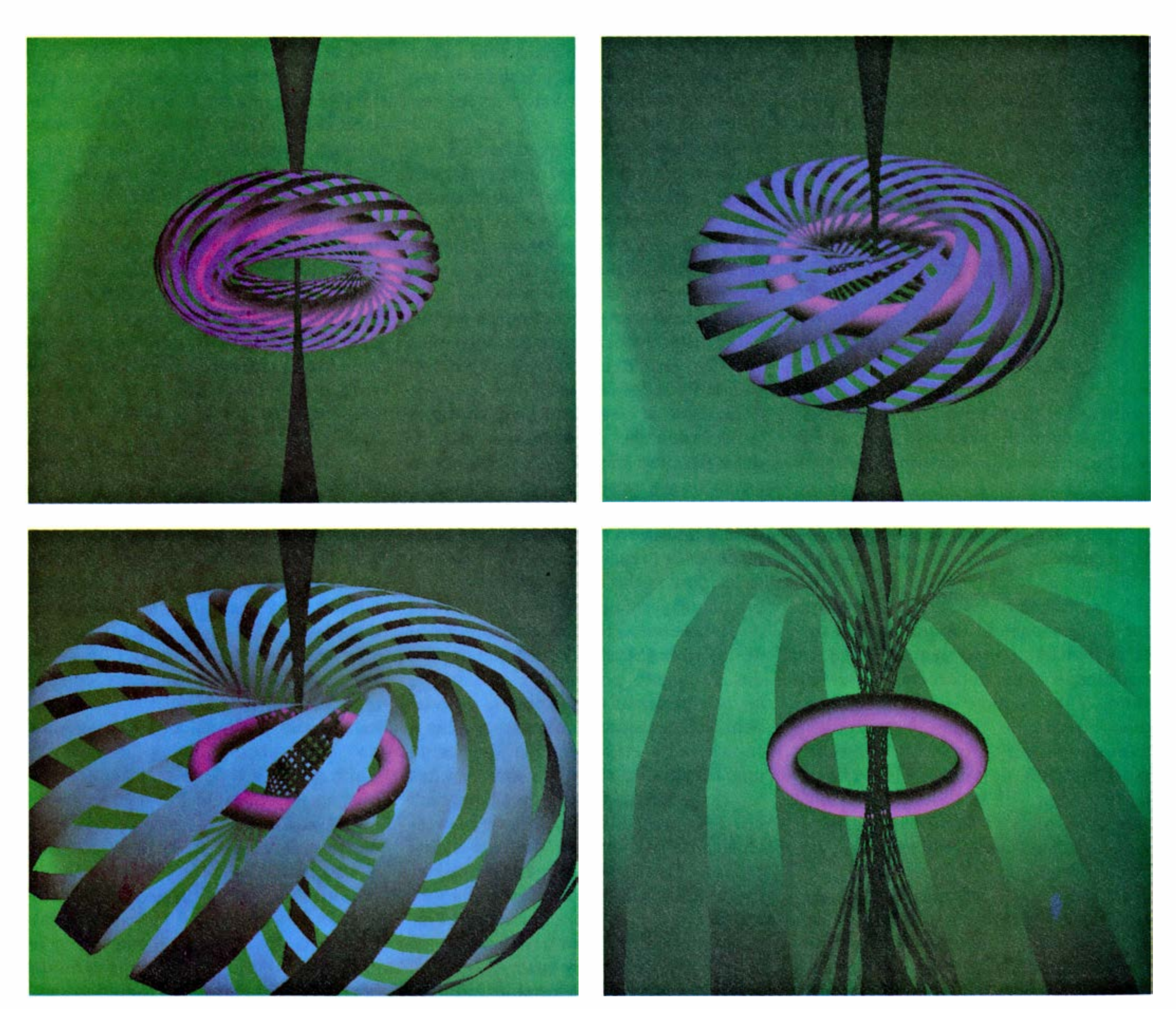
A cada uno de esos seis planos le corresponde un determinado tipo de rotación, especificada mediante una matriz de 4 por 4, compuesta por 16 números. Las seis matrices de giro se denominan *rot12*, *rot13*, *rot14*, *rot23*, *rot24* y *rot34*. El usuario del programa tiene que dar el nombre del tipo de matriz elegida y el ángulo del giro que la matriz va a generar. Por ejemplo, al escribir "rot23" y a continuación "60", el efecto sería el de un giro de 60 grados en torno al plano determinado por los ejes segundo y tercero.

Supongamos que se desee confinar la rotación del hipercubo a las dimensiones tercera y cuarta, que corresponden al giro más misterioso de todos. Se aplica la matriz de giro *rot34*. Los elementos de esta matriz son ceros, unos y otros tres números, distribuidos de la forma siguiente:

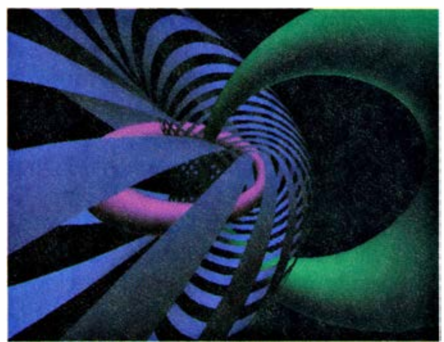
$$\begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & -b & a \end{matrix}$$

Se selecciona el ángulo de giro deseado, en grados, que se almacena en la variable *ang*. Los números a y b son función de *ang*: a es igual a *cos*(*ang*) y b es igual a *sen*(*ang*), siendo *cos* y *sen* las funciones trigonométricas seno y coseno.

La regla para generar las cinco ma-



4. Sucesión de toros concéntricos, análogos de los paralelos de la esfera proyectados desde la hiperesfera en el espacio de tres dimensiones


5. Proyección del movimiento de los toros al girar la hiperesfera, análoga a la proyección del movimiento de los paralelos correspondientes a la rotación de la esfera

trices de giro es sencilla. Las *a* aparecen siempre en la diagonal principal de cada una de esas matrices, siendo las filas que ocupan las correspondientes a las dos dimensiones afectadas por la rotación. Las *b* figuran en todas las demás intersecciones de filas y columnas correspondientes a las dimensiones que han de girar. Los demás números de la diagonal principal son unos; todos los restantes elementos de la matriz son ceros. Por ejemplo, *rot13* es la matriz siguiente:

$$\begin{matrix} a & 0 & b & 0 \\ 0 & 1 & 0 & 0 \\ -b & 0 & a & 0 \\ 0 & 0 & 0 & 1 \end{matrix}$$

Una vez elegida la matriz de giro deseada, HIPERCUBO se la asigna a una matriz especial, *rotar*. La asignación se hace cómodamente mediante un bucle doble; el bucle interior correspondería a la sucesión de números de cada fila y, el exterior, a la sucesión de filas.

El cálculo de las coordenadas del hipercubo, tras el giro, se efectúa "multiplicando" la matriz *cubo* por la matriz *rotar* que se haya elegido. El producto de ambas matrices queda temporalmente almacenado en una tercera matriz, llamada *temp*, que contiene las coordenadas del hipercubo recién girado. Tal producto se determina de acuerdo con las reglas de multiplicación de matrices, operación habitual y perfectamente definida en las matrices de números. El resultado se alcanza tras una ordenada orgía de multiplicaciones y sumas, integradas en tres bucles encajados.

Al igual que *vert* y que *cubo*, la matriz *temp* es de 16 por 4, formada por 64 números; contiene las cuatro coordenadas correspondientes a cada uno de los 16 vértices del hipercubo girado. Cada número *temp*(*i*, *j*) de que consta la matriz está designado por un par de índices, *i* y *j*. Por ejemplo, *temp*(*13*, *3*) es la tercera coordenada del décimotercer vértice del hipercubo girado. Ese valor es la suma de cuatro productos de números tomados de las matrices *cubo* y *rotar* según una regla perfectamente determinada [*véase la figura 6*]. Por consiguiente, el bucle más interno del programa puede controlarlo un índice *k*, que vaya de 1 a 4; la salida de ese bucle proporciona el valor de un elemento de *temp*.

Seguidamente se aloja el bucle *k*, un bucle intermedio, cuyo índice es *j*. El bucle *j* calcula las cuatro coordenadas del *i*-ésimo vértice, según el procedimiento que acabo de esbozar; dicho de otro modo, lo que hace es rellenar los cuatro lugares de contenido de la *i*-ésima fila de *temp*. Finalmente, el bucle *j* se sitúa en el interior del bucle más exterior, que tiene índice *i* y que calcula la totalidad de las 16 filas de *temp*; al finalizar su ejecución, *temp* contiene las coordenadas de todos los vértices del hipercubo girado, que resplandece en su nueva posición. Para mostrarlo sobre la pantalla del ordenador se requiere un nuevo bucle doble, que remplace las antiguas coordenadas de posición contenidas en *cubo* por las recién calculadas, contenidas en *temp*.

Un hipercubo tiene cuatro dimensiones, pero la pantalla del ordenador sólo dos. Conviene, por tanto, estipular que las dos primeras dimensiones, o coordenadas, del hipercubo correspondan a coordenadas en la pantalla. La forma más sencilla de tratar las dimensiones tercera y cuarta consiste en despreciarlas. Eso es justamente lo que hace la técnica de presentación que voy a describir, si bien es posible perfeccionarla, con lo que el objeto resultante quedará proyectado en su casi demoniaca complejidad, y hacer algo más visibles las coordenadas tercera y cuarta.

Para presentar una versión esquelética del hipercubo el programa no tiene más que mostrar sus aristas. Dado que el hipercubo tiene 32, basta con que la sección destinada a presentación en HIPERCUBO trace las líneas adecuadas. Pero, ¿en qué orden? El número de posibilidades es casi infinito, y por ello la respuesta sea seguramente cuestión de elección personal y de sentido estético. A pesar de todo, resulta difícil resistirse a trazar con las aristas un sendero de Euler, así llamado en honor del matemático Leonhard Euler. Para que un lápiz dibuje sendas de ese tipo es preciso no levantarlo del papel y no recorrer ningún tramo más de una vez. Las aristas consecutivas de un hipercubo, visto como un sendero euleriano, tienen un vértice común.

En torno a él, a su través, arriba y abajo corre veloz la senda euleriana al ir conectando los vértices. He aquí una, que me ha dado la impresión de ser muy bella, y que se obtiene al unir los vértices del hipercubo en la siguiente secuencia: 0, 1, 3, 2, 6, 14, 10, 8, 9, 11, 3, 7, 15, 14, 12, 13, 9, 1, 5, 7, 6, 4, 12, 8, 0, 4, 5, 13, 15, 11, 10, 2, 0. Estos vértices están almacenados en una tabla llamada *senda*, cuyo subíndice será *i*; el *i*-ésimo vértice de esta serie de 33 vértices se designa *senda*(*i*). Existen, para cada valor de *i*, instrucciones para consultar la primera y segunda coordenadas tanto de *senda*(*i*) como de *senda*(*i*+*1*). Seguidamente, a fin de conectar los dos puntos debe invocarse el mandato preciso para el trazado de rectas de que disponga nuestro lenguaje de programación. Tanto la consulta como el trazado de la recta se encuentran en un único bucle, cuyo índice es *i*, que se ocupa de ir trazando una línea desde cada vértice de la sucesión hasta el siguiente.

Entremos ahora en las complicaciones de carácter visual (y psicológico). Dos son los procedimientos típicos para presentar la tercera dimensión del hipercubo. El método ortográfico se limita a despreciar sin más la tercera dimensión, proyectándose todos los vér-

tices sobre la superficie plana de la pantalla de visualización sin tener en cuenta lo muy por detrás de ella que puedan encontrarse. Si se hace la proyección desde un único punto (proyección estereográfica), los vértices aparecen sobre la pantalla como si se tratara de sombras lanzadas por un foco luminoso puntual, situado sobre el centro de la pantalla, y a una cierta distancia por detrás del hipercubo. Aunque la observación de las sombras equivale a una observación del hipercubo desde atrás, tal observación resulta visualmente indistinguible de la observación frontal.

Para lograr en HIPERCUBO el efecto de una proyección estereográfica se supone que la tercera coordenada de cada vértice es igual a la distancia entre el vértice y la pantalla de presentación, medida en la dirección del punto imaginario fuente de luz. El programa calcula la razón de semejanza de triángulos proporcionales y determina así el coeficiente necesario para convertir las dos primeras coordenadas de los vértices en coordenadas de puntos en pantalla. Por ejemplo, si la imaginaria fuente de luz se encontrara a 20 unidades por detrás de la pantalla, un vértice situado en (5, −7, 11, 8) podría proyectarse en ella multiplicando por 20 cada una de las dos primeras coordenadas, y dividiendo cada resultado por 20 −11, esto es, 9.

Temía yo tener que incluir en el poco espacio que resta una descripción completa del proceso de creación de imágenes estereoscópicas que permitieran atisbar la cuarta dimensión del hipercubo. Existe una técnica general para crear imágenes estereoscópicas, a la que confío dedicar un artículo en el futuro. Sin embargo, para el programa del hipercubo, Banchoff y sus colegas han adoptado un procedimiento mucho más sencillo. Para cada una de las posiciones del hipercubo, crean un nuevo par de imágenes, aplicando *rot14* con un ángulo de 3 grados en un sentido y 3 grados en el otro. La dimensión 1 es la dirección paralela a la alineación horizontal definida por los ojos del observador, y la dimensión 4 es el objeto del ejercicio. Los dos giros anteriores describen muy bien las dos vistas del hipercubo desde los ojos del observador: basta imaginar que las dos visuales convergen en el centro del hipercubo formando un ángulo de 6 grados.

Los lectores que deseen captar las emociones de las películas en "tres dimensiones" pueden hacerse unas gafas para visión estereoscópica con celofán de dos colores, azul y rojo. En tal caso se hace funcionar dos veces el programa HIPERCUBO, una para cada uno de los dos pequeños giros. El resultado de la primera rotación ha de dibujarlo el programa en color azul y, el de la segunda, en color rojo. Si al construir las gafas se las hace reversibles, no es preciso preocuparse de qué color corresponde a cuál.

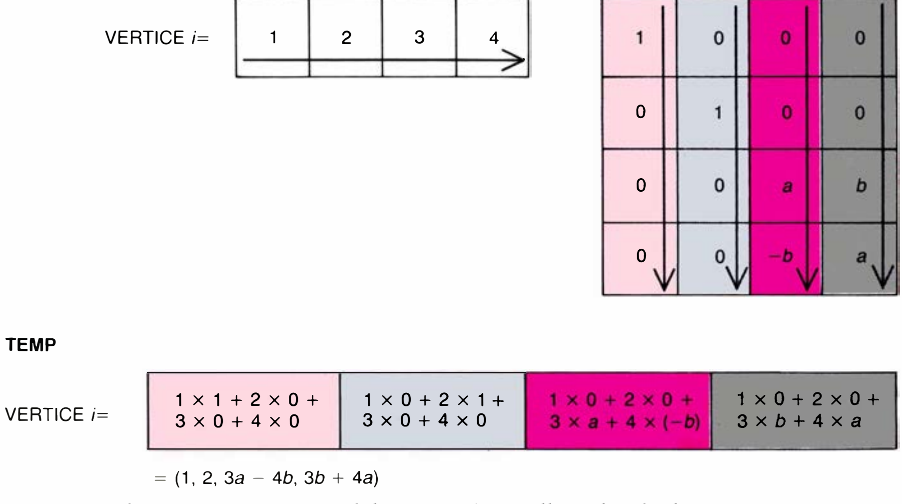
Por mi parte, no queriendo luchar con el papel de celofán, he preferido aprender a fundir en una las imágenes estereoscópicas por pura fuerza de voluntad. La técnica exige reducir de tamaño las dos imágenes resultantes de los pequeños giros y trasladarlas a posiciones de la pantalla horizontalmente contiguas, y no traslapadas. Conviene que sean del mismo color, por lo que basta con un monitor monocromático, y su separación no debe ser mayor que la de los ojos del observador. No hay que mirar fijamente a las imágenes; por el contrario, debe mirarse a un punto situado entre ambas, e infinitamente más allá. Los dos hipercubos parecerán deslizarse y zangolotear fundiéndose ambos en uno.

Aun cuando las dimensiones tercera y cuarta no reciban tratamiento especial, HIPERCUBO puede engendrar imágenes muy similares a las mostradas en la secuencia gráfica de Banchoff. Merced a una serie de giros de ángulos pequeños, los lectores pueden ver cómo los dos burdos toros se hinchan, se deshinchan y se regeneran, de modo muy similar a como lo hacen sus primos, más lisos, de la ilustración de la hiperesfera giratoria.

Evidentemente, había sido el programa HIPERCUBO el responsable de la desaparición de mi amigo Magi. El final feliz de esta chifladura tetradimensional llegó con una llamada telefónica. Nada sorprendentemente, me contó cosas estupefacientes. "Seguramente creerás que estoy loco." (La frase siempre le pone a uno sobre aviso.) "He estado flotanto hasta ahora por la cuarta dimensión. Vi pasar a mi lado una sección transversal de mi casa. Entonces, me acerqué y le hice cosquillas a mi gato en los riñones..."

Ahorraré a los lectores otros penosos detalles de la conversación. Baste decir que he logrado convencer a Magi para que nunca más vuelva a hacer funcionar el HIPERCUBO, y que mantenga toda ulterior exploración exclusivamente en el plano intelectual. Me asegura haber seguido mis consejos, y ahora mantiene haber realizado muchos descubrimientos maravillosos, merced a su mayor capacidad de comprensión y de visión interior, artificialmente amplificada. Por ejemplo, me ha venido con un par de enigmáticos problemas que a mi parecer vale la pena presentar aquí a los lectores.

Pensemos por un momento en la siguiente sucesión de objetos: un segmento unitario, un cuadrado unitario, un cubo unitario, y así sucesivamente. El n-ésimo elemento de la sucesión es el análogo n-dimensional del cubo. Efectuemos ahora sobre los objetos dos experimentos conceptuales: trazar la diagonal del cubo n-dimensional e inscribir una esfera n-dimensional en el interior del cubo n-dimensional. La diagonal se extiende desde un vértice hasta el opuesto más alejado. ¿Qué le ocurre a su longitud conforme n va haciéndose cada vez mayor? Asimismo, ¿qué le sucede al hipervolumen de la esfera n-dimensional al crecer n indefinidamente? Las respuestas de Magi



6. Modo de girar un vértice del hipercubo y llevarlo de la tercera dimensión a la cuarta

mal pueden parecer cuerdas; las expondré en el artículo del mes próximo.

En mi artículo de marzo describí dos programas, CUMULO y SUPERCUMULO, que simulan la evolución de un cúmulo estelar. Me alegra que en unos cuantos miles de hogares esos programas hayan proporcionado una nueva forma de entretenimiento que, al menos temporalmente, ha dejado un poco de lado la televisión. Sin duda algunos de esos universos de sala de estar evolucionan como debieran, pero tal vez otros estén creando problemas. No hemos de buscar la causa en nuestras estrellas, sino en nosotros mismos.

Brian Davis, de Ann Arbor, Michigan, y Peter Fortescue, de La Jolla, California, tropezaron en SUPERCUMULO a causa de las ecuaciones de la aceleración. Según entiendo, tales dificultades se resuelven sustituyendo la fuerza f de la ecuación de la página 96 por la aceleración a debida a esa fuerza. Para calcular a, se divide f entre la masa de la estrella atraída. Andrew M. Odlyzko, de los Laboratorios AT & T Bell, señaló que las coordenadas de posición de la tabla de la página 97 vienen dadas por múltiplos de 1000 unidades astronómicas (U. A.), y no directamente en U. A. Nuestro universo se desplegará ahora correctamente.

A veces, cuando se observa el vivaz funcionamiento de CUMULO en un monitor, resulta difícil advertir qué estrellas se encuentran en primer término y cuáles más atrás. Albert C. English, de Delray Beach, Florida, y Peter Sterans, de Lodi, California, han preparado programas de presentación especiales, que generan dos imágenes contiguas de los cúmulos; una, tal como la veríamos con el ojo derecho y, la otra, con el izquierdo. Los lectores capaces de resolver las pegas de la visión estereoscópica podrán observar los cúmulos, lo mismo que el hipercubo, con sorprendente profundidad y verismo.

Varios lectores habían escrito ya programas similares a SUPERCUMULO, pero los habían aplicado a nuestro propio sistema solar. La misma aplicación sería posible con SUPERCUMULO. Los más emprendedores pueden consultar las masas, posiciones y velocidades de los 10 principales cuerpos correspondientes a un tiempo de referencia. Podemos arreglárnoslas para ver la evolución del sistema solar desde arriba: basta esperar unos cuantos minutos para llegar al año 2000. Geoffrey L. Phillips, de St. Louis, Missouri, preparó una simulación para el sistema Tierra-Luna que incluye un pequeño vehículo espacial, desprovisto de masa. No es hazaña trivial lanzar la nave desde la Tierra y lograr que entre en órbita lunar. Los más experimentados podrían ensayar el lanzamiento de un vehículo Voyager a un gran periplo en torno a los gigantes gaseosos que acabe cuando la nave abandone el sistema solar.

William A. Hoff, de Champaign, Illinois, determinó dinámicamente los incrementos de tiempo utilizados en la simulación, situando al principio del programa una variable llamada *dvmax*. En el curso de los cálculos del movimiento estelar, el programa va hallando siempre la aceleración máxima, *amax*, de una estrella. El incremento de tiempo correspondiente al paso siguiente es *dvmax* dividida entre *amax*. Esta técnica impide que ninguna velocidad supere *dvmax*.

En el artículo del mes pasado propuse yo un minitest del C.I. y dejé pendientes algunas cuestiones relativas a sucesiones numéricas. El primer problema del minitest es buen ejemplo de la ambigüedad típica de tales problemas. Consistía en completar la sucesión 3, 7, 16, 35,... Si a cada término se le resta el doble del anterior, se obtiene la sucesión 1, 2, 3,..., que sería la segunda fila de la pirámide. Según este razonamiento, el término que falta sería el doble de 35, más 4, o sea, 74. Por otra parte, si se construye una sencilla pirámide de diferencias de tres filas, la tercera fila da la sucesión 5, 10, y parece razonable continuar tal sucesión con 15. El término que añadir debería ser 69; los programas descritos hubieran pasado por alto esta solución. He aquí las respuestas del test: la letra que falta es la H; la palabra que añadir es "por"; la palabra excepcional es "identidad"; colocadas en orden sus letras, el nombre de ciudad que no está en Italia es Madrid, y la analogía visual corecta es la número 2.

Las sucesiones numéricas de la página 96 se completan con 350 y 22, respectivamente. La primera sucesión de la figura 1 puede resolverse aplicando una regla generalizada de diferencias, con $k = 3$, y después una regla generalizada de cocientes; el término que falta es 324. La segunda sucesión debería derrotar a todos quienes no hayan probado a escribir SE Q, exceptuados, posiblemente, los más pacientes y tenaces aficionados a rompecabezas. Puede resolverse mediante dos reglas de tipo cociente; el valor de k en la primera es 5. El número que falta es −65.551.

Taller y laboratorio

Maravillas del retro-reflector, espejo que elimina la distorsión del haz luminoso

Jearl Walker

Un espejo ordinario refleja un haz luminoso hacia la fuente de origen en el caso exclusivo de que esté orientado perpendicularmente al haz. Por contra, un retro-reflector retorna los rayos a la fuente cualquiera que sea su ángulo de incidencia. Hay un experimento, nuevo e ingenioso, que puede hacerse con un conjunto de pequeños retro-reflectores: si un haz es distorsionado por una hoja de plástico corrugado, pongamos por caso, antes de que llegue al dispositivo, éste suprime la distorsión invirtiendo el camino de los rayos a través del plástico. David M. Pepper, de Hughes Research Laboratory, ha redactado un manuscrito, que me sirvió de base para este artículo, donde se describen este y otros experimentos relativos al fenómeno. Tales experimentos se basaron sobre todo en las investigaciones de Harrison H. Barrett y Stephen F. Jacobs, de la Universidad de Arizona.

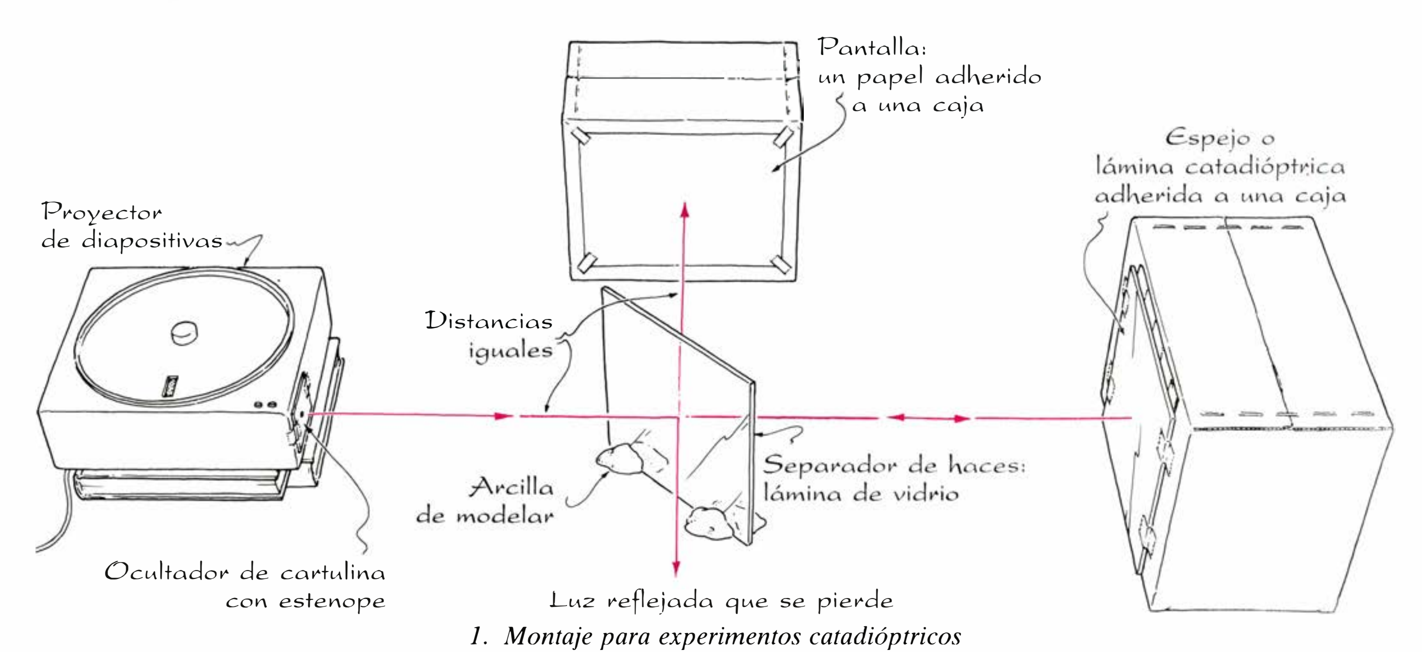
Pueden conseguirse dispositivos retro-reflectores en las tiendas de artículos deportivos, bajo la forma de hojas de plástico flexible, y en las de accesorios de automóviles como catafaros. Cosidos a la ropa, montados en una bicicleta, encolados a las rayas de las carreteras o unidos a una variedad de otros objetos, estos reflectores catadióptricos, que así se llaman también, hacen que los objetos se vean mejor de noche. Por ejemplo, cuando las luces de un automóvil inciden en un catafaros, u otro dispositivo similar, la luz regresa con gran intensidad y sirve de advertencia al conductor. En realidad, el éxito práctico de estos sistemas se debe a una cierta imperfección de los mismos. En efecto, si un catafaros reflejase los haces luminosos de manera muy precisa y no los dispersara un poco, éstos regresarían a los faros del automóvil y no hacia el conductor.

Para efectuar los experimentos de Pepper debe adquirirse una hoja cuadrada de plástico catadióptrico flexible, de tres centímetros de lado como mínimo. Si se encontraran sólo en bandas más estrechas, pueden cortarse en pedazos y juntarlas para formar el cuadrado, que hay que fijar a una caja rígida con cinta adhesiva. Se coloca un proyector de diapositivas de forma que ilumine el plástico y, si es posible, se retira la lente frontal. En la parte anterior del proyector se monta un trozo de cartulina gruesa en la que se ha practicado un orificio de dos o tres milímetros de diámetro. Este pequeño orificio servirá de fuente luminosa puntual. Si no se dispone de proyector, puede recurrirse a una linterna potente, e incluso a una bombilla de incandescencia ordinaria, que proyecte su luz a través del estenope.

Al objeto de observar la forma del haz de retorno, debe separarse parte de éste para dirigirlo hacia una pantalla. Esto puede hacerse colocando una lámina de vidrio en el camino del haz, de modo que forme un ángulo de 45 grados con el mismo. Este vidrio debe estar bien limpio, pero no hace falta que sea de una calidad óptica sobresaliente. Puede servir el cristal de un portarretratos. El vidrio puede fijarse en su posición con arcilla de modelar, o bien con el marco del portarretratos al que se habrá retirado el cartón del reverso.

Cuando el haz procedente del estenope llega al vidrio, se separa en dos. Uno de los haces prosigue hacia el catadrióptico y el otro se pierde en la habitación. El primero se refleja en el plástico y regresa al vidrio, donde éste lo vuelve a dividir en dos. De éstos, uno atraviesa el vidrio y retorna al estenope y el otro se refleja hacia una hoja de papel que hace las veces de pantalla de observación. Como la luz se refleja tanto en la superficie frontal como trasera del vidrio, la pantalla exhibe dos imágenes ligeramente desplazadas. Si se usa como separador de haz un vidrio más delgado resulta un desplazamiento menor y, por tanto, una imagen más nítida.

El separador de haz debe colocarse equidistante de la pantalla y del estenope, para que la luz devuelta por el plástico recorra la misma distancia hasta cada uno de ellos. Cuando todo



1. Montaje para experimentos catadióptricos

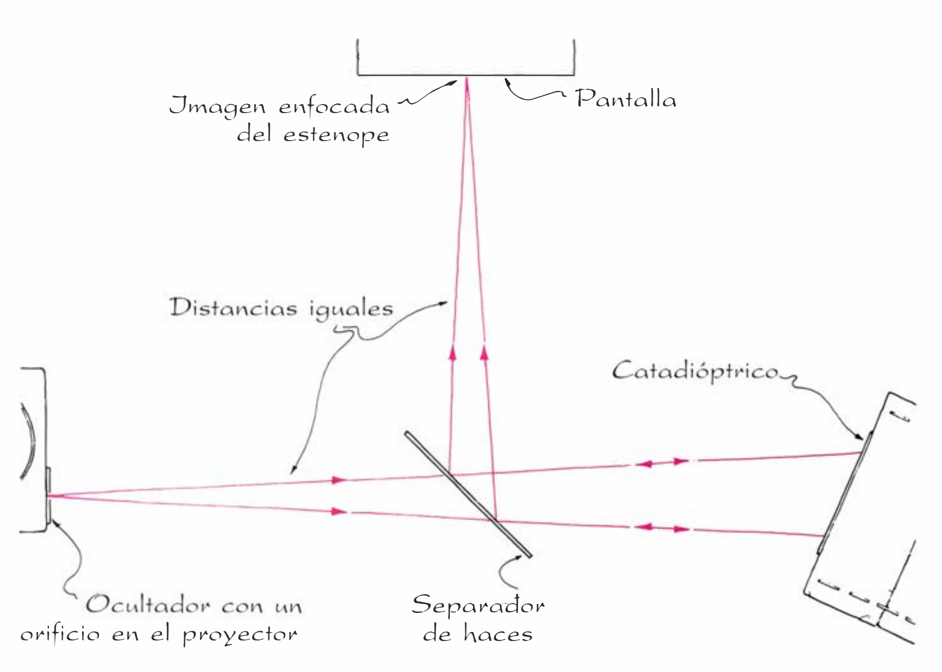
se halle dispuesto, sobre la pantalla aparecerá un círculo pequeño y brillante, el cual no es sino una imagen del estenope producida por la luz que hubiera retornado al mismo de no haber mediado la acción del separador. Esa imagen tiene aproximadamente el mismo tamaño que el estenope, pero con bordes borrosos a causa de la doble reflexión en el vidrio y las imperfecciones del catadióptrico.

Mirando a la pantalla a través del separador, se ve una manchita brillante que parece alojada en aquélla, y que puede fácilmente confundirse con la imagen devuelta por el plástico. No es tal: se trata de una imagen formada por el ojo del observador con la luz reflejada hacia él por el separador. Tales imágenes, llamadas imágenes virtuales, son muy corrientes en los espejos y, en nuestro experimento, el separador actúa como un espejo débil o semitransparente. Mirando desde una visual apropiada, es posible alinear la imagen virtual con la imagen real producida en la pantalla. Para eludir la complicación de la imagen virtual, lo mejor es no mirar por el separador.

Variando la distancia entre el catadióptrico y el separador de haz, la imagen del estenope en la pantalla no se mueve, ni se altera tampoco si movemos el catadióptrico de un lado a otro del haz. Girándolo respecto al haz, la imagen no cambia hasta que el ángulo es grande. En resumen: el catadióptrico devuelve los haces luminosos a la fuente aunque no esté colocado perpendicularmente a los haces.

Pongamos un espejo en vez del plástico. La pantalla entera se nos ofrecerá bañada en una luz tenue y no se verá la imagen del estenope. Probemos con un espejo cóncavo: éste dará en la pantalla una imagen nítida del estenope únicamente si su foco coincide con el orificio. Si movemos el espejo a lo largo del haz, o lo giramos respecto al mismo, la imagen se ensanchará, distorsionándose, y acabará por desaparecer.

El efecto de estos espejos y del catadióptrico puede comprenderse mejor con ayuda de la figura 4. En la parte *a* se ilustra la reflexión producida por un espejo plano. El estenope está representado por un punto del que emanan ondas luminosas. Estas ondas exhiben cada una un frente de onda que forma un casquete esférico. Uno de esos frentes se muestra en sección transversal como parte de un círculo con centro en la fuente. El radio de ese círculo aumenta a medida que el frente de onda se aleja de la fuente. Conforme el frente llega al espejo, se va reflejando

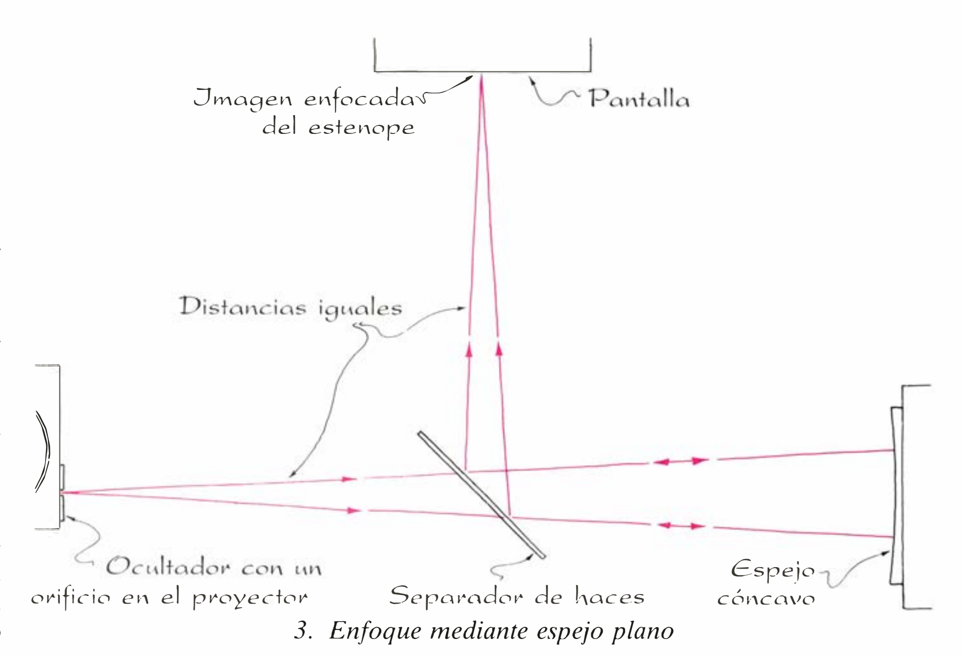


2. Vista en planta del enfoque en un catadióptrico

y genera un nuevo frente que se propaga en sentido contrario. El frente reflejado es parte de un círculo centrado en una fuente puntual virtual situada a la derecha del espejo y a la misma distancia de éste que la fuente real. El frente reflejado no puede formar una imagen del estenope porque crece indefinidamente, sin llegar jamás a focalizarse. Por consiguiente, cuando el separador dirige parte de la luz a la pantalla, la totalidad de ésta queda iluminada con el frente en expansión.

En la parte *b* se explica el efecto del espejo cóncavo. Este refleja y enfoca el frente incidente, pero la luz sólo regresará a la fuente si el estenope se halla en el foco del espejo. Como la pantalla está a la misma distancia del separador que el estenope, sobre aquélla se forma una imagen nítida de éste. Para otras posiciones del espejo, la luz llega a la pantalla más esparcida. Y si el espejo no está orientado simétricamente en torno al haz, el enfoque de la luz se sesga y, como resultado, la imagen se deforma o desaparece.

En la parte *c* se ilustra el efecto de un sistema especular muy especializado afín a los catadióptricos comerciales. Recibe el nombre de espejo conjugador de fase (véanse "Conjugación de fase óptica", de Vladimir V. Shkunov y Boris Ya. Zel'dovich, *Investigación y Ciencia*, febrero de 1986, y "Aplicaciones de la conjugación de fase óptica", de David M. Pepper, *Investigación y Ciencia*, marzo de 1986). Este catadióptrico ideal invierte exactamente la dirección de los rayos luminosos que



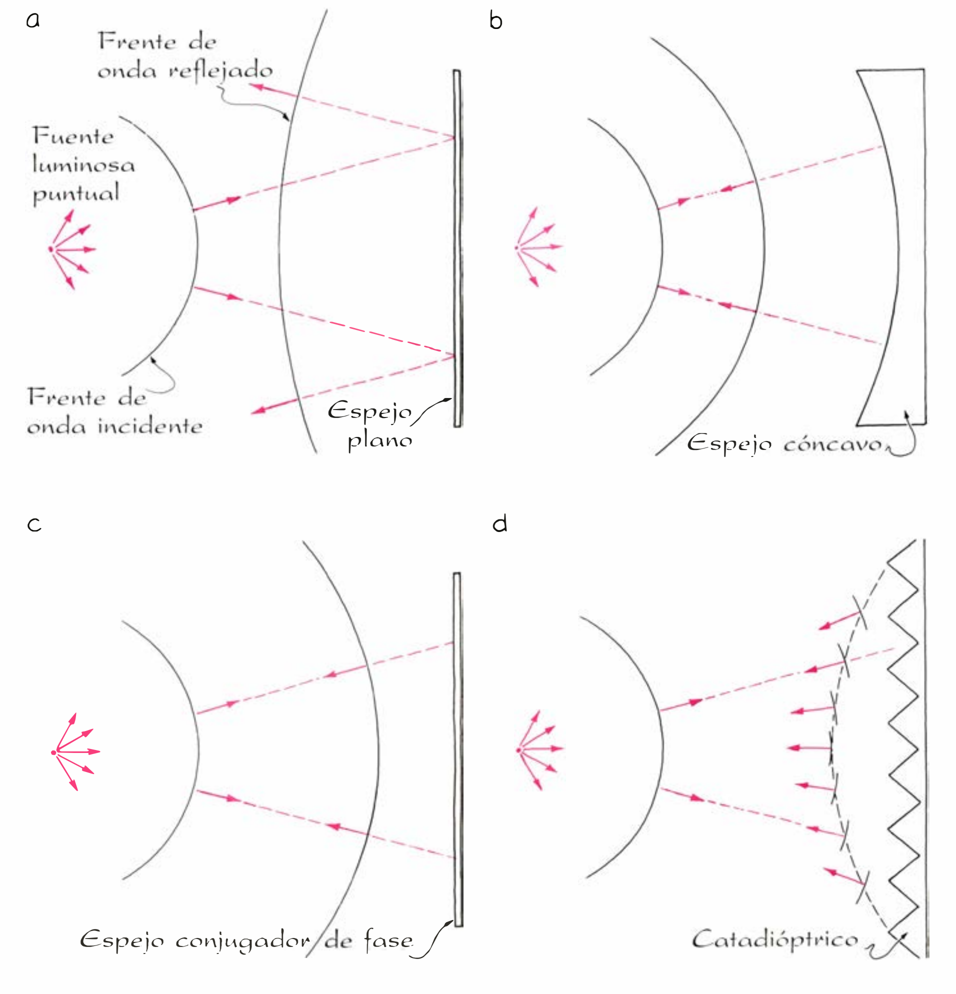
3. Enfoque mediante espejo plano

inciden en él. Más aún, el frente reflejado conserva la forma y orientación del frente incidente. Con este espejo ideal, toda la luz devuelta que atraviese el separador pasará por el estenope, y la fracción que vaya a la pantalla formará una imagen exacta del mismo.

Pero el plástico de Pepper no es, ni mucho menos, un retro-reflector ideal. Cada catadióptrico de una lámina devuelve una parte del frente, preservando su forma y orientación, tal como se muestra en la parte *d* de la figura. Porque los catadióptricos yacen en un plano, las distintas secciones no encajan unas con otras para restaurar el frente inicial. Antes bien, se disponen a lo largo de una curva idéntica al frente reflejado por el espejo plano. Por esta causa, se dice de las láminas que son pseudoconjugadoras. Cada sección del frente regresa al estenope, o a la imagen de éste en la pantalla, pero su reunión carece de la coordinación que se observaría con un espejo de conjugación de fase. Aun así, en la pantalla se forma una imagen borrosa del estenope conforme a ella van llegando las secciones del frente.

Pueden encontrarse láminas de plástico retro-reflectoras de dos tipos. Uno de ellos está constituido por prismas incrustados en el plástico. Un rayo incidente se reflejará en una cara interna del prisma, luego en una segunda y posiblemente en una tercera. Así que saldrá del plástico paralelamente al rayo incidente. Este retro-reflector no es perfecto, ya que el rayo devuelto está desplazado con respecto al incidente y los prismas invierten en la reflexión cada sección del frente de onda. (Dada la pequeñez del retro-reflector, la porción de frente puede considerarse plana.) En ciertos catadióptricos comerciales, los prismas están agrupados en células hexagonales de 0,15 milímetros de diámetro. Las zonas comprendidas entre las células no intervienen en el fenómeno.

El Scotchlite, que fabrica 3M Corporation, consiste en cuentas de vidrio inmersas en un soporte flexible. Cuando llega un rayo, se refracta hacia el fondo de la cuenta, donde se refleja y vuelve hacia adelante. Así, al abandonar la cuenta, es paralelo al rayo incidente. Este retro-reflector es imperfecto porque desplaza el rayo resultante e invierte el frente de onda.



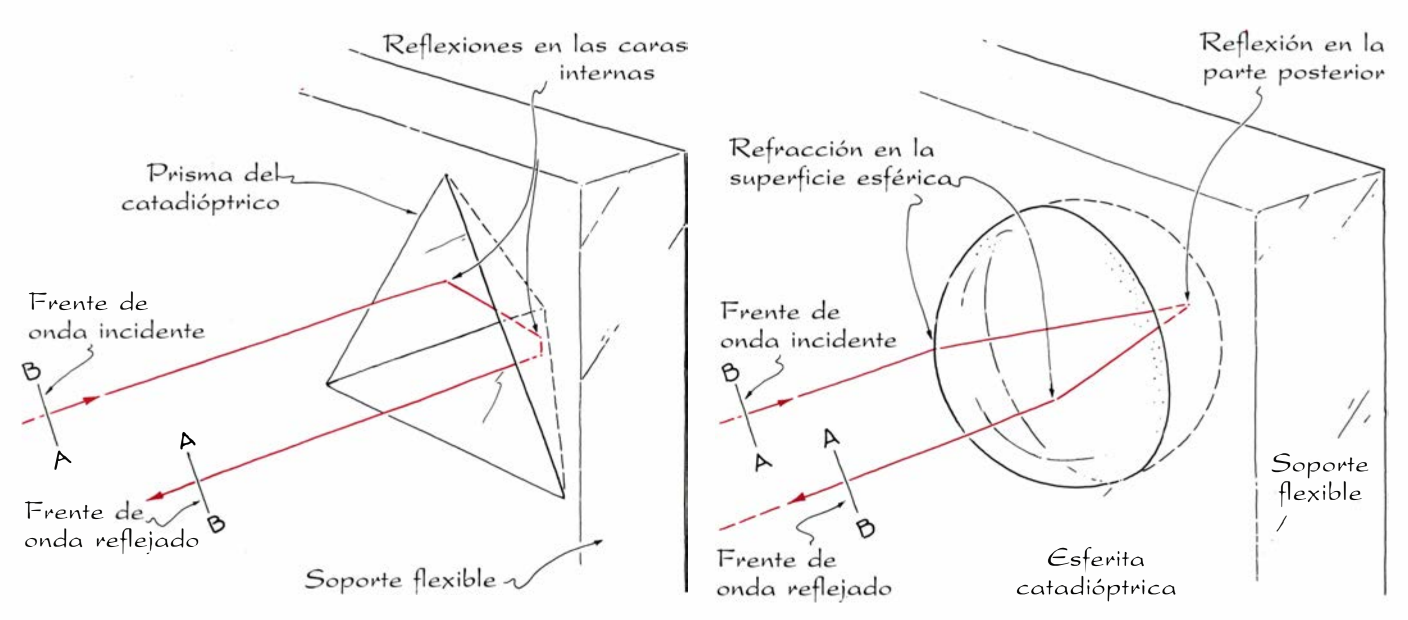
4. Reflexiones en diferentes tipos de espejos

Hay una nueva aplicación de estos sistemas, original e ingeniosa, en la que se elimina la distorsión sufrida por un haz luminoso al atravesar un material no uniforme. En el trayecto de la luz procedente del estenope se coloca un espejo cóncavo, situándolo de modo que en la pantalla se consiga la mejor imagen posible. Ante el espejo se intercala un elemento distorsionador. Este podría ser uno de esos plásticos rizados que sirven de cubierta a las lámparas fluorescentes, una platina de vidrio cubierta de cola de aviación, un trozo de plástico "de burbujas" para embalajes o cualquier otro material no uniforme y transparente. El elemento destruye la imagen del estenope, dispersando la luz y llenando la pantalla con una intrincada configuración luminosa. Desplazando el elemento de un lado a otro de la abertura del espejo, aquella configuración varía caóticamente.

Sustituyamos ahora el espejo por un sistema catadióptrico. De éste no importa su posición exacta ni su orientación, pero el elemento distorsionador debe hallarse directamente ante él, y ambos deben estar iluminados. Aunque el elemento distorsionador entremezcle las trayectorias de los rayos y parezca echar a perder toda información acerca de la forma del estenope, en la pantalla aparecerá una imagen de éste. No será perfecta, sino tenue y borrosa.

La aparición de la imagen se debe a que el catadióptrico invierte casi exactamente los rayos que emergen del elemento distorsionador. Los rayos, por tanto, regresan casi a través de los mismos accidentes deformadores que en su viaje de ida. Ese segundo paso invierte la distorsión de cada rayo, se reconstruye el haz y en la pantalla se forma la imagen del estenope. Pepper sugiere considerar que el catadióptrico produce una inversión temporal en los rayos. Moviendo el elemento por delante del catadióptrico, la imagen del estenope cambia muy poco o nada. ¿Por qué? Pues debido a que el tiempo que tarda la luz en ir y volver desde una irregularidad del elemento hasta el catadióptrico es tan breve que la irregularidad permanece inmóvil durante el recorrido.

Adviértase que la distorsión se elimina porque la luz atraviesa dos veces el elemento distorsionador. (Si éste se sitúa delante de la pantalla o del estenope, la luz lo atravesará sólo una vez.) Los rayos deben hacer su segunda travesía por el elemento siguiendo aproximadamente los mismos trayectos que en la primera. Recuérdese que los



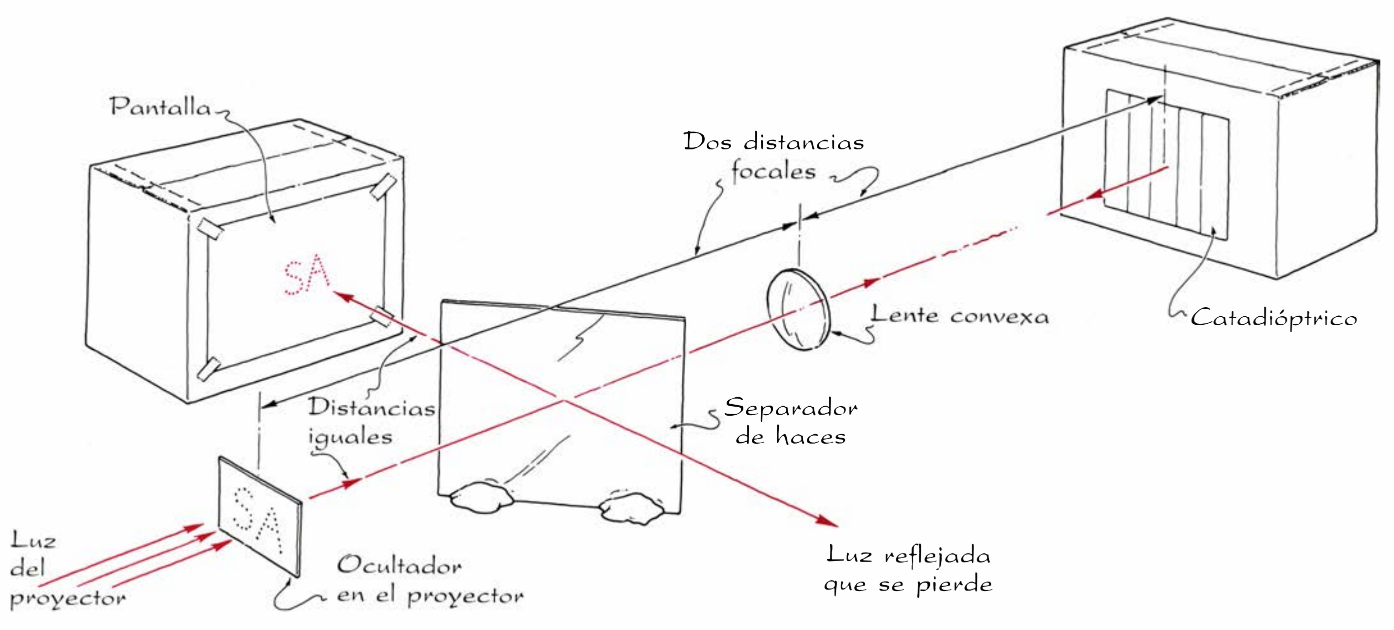
5. Dos tipos de catadióptricos

retro-reflectores retornan unos rayos levemente desplazados con respecto a los incidentes. Si el elemento y el catadióptrico estuvieran excesivamente separados, los rayos devueltos encontrarían unos accidentes diferentes en su segundo viaje y no desaparecería la distorsión. Una separación grande podría también implicar que el catadióptrico interceptase y devolviese menos luz, con lo que resultaría una imagen más tenue.

Unas experiencias similares pueden hacerse remplazando el proyector y estenope por un láser. Pero aquí la imagen aparece desfigurada por unas interferencias suplementarias. Esas interferencias dan lugar a una complicada configuración de líneas claras y oscuras, provocadas, en parte, por la difracción de la luz en las motas de polvo existentes en el separador y en las irregularidades del vidrio. Parte de la luz que llega a la pantalla se ha deformado sólo una vez y conserva por ello la distorsión. Otra causa de interferencia en la luz que llega a la pantalla reside en la periodicidad de las cuentas o de las células hexagonales de las láminas retro-reflectoras.

Al hacer este ensayo puede sorprendernos otro punto luminoso que aparece en la pantalla. A causa de la constitución imperfecta del catadióptrico, parte de la luz que devuelve la refleja del mismo modo que un espejo plano. Si el plano del catadióptrico es perpendicular al haz, esa reflexión especular llega a la pantalla y puede confundirse con la imagen del estenope formada por retro-reflexión. El punto adicional se identifica fácilmente porque se mueve al inclinar el catadióptrico con respecto al haz. Puede eliminarse poniendo una cartulina entre el catadióptrico y la pantalla, de manera que el haz de láser pase rasando el borde de la cartulina. Luego se gira el catadióptrico hasta que la reflexión especular aparezca en la cartulina y se le impida su marcha hacia la pantalla.

Con un láser, un sistema catadióptrico y lentes y espejos ordinarios pueden hacerse numerosos experimentos. Me divierte especialmente lanzar humo o polvo fino en el trayecto del haz de láser y seguir su recorrido. En una ocasión hice pasar el haz por el separador, luego se reflejaba en un espejo plano inclinado y lo recuperaba después de pasarlo por el catadióptrico. Con ese recorrido se ponen de manifiesto dos tipos de reflexión diferentes: la ya co-



6. Creación de una imagen a partir de un grupo de estenopes

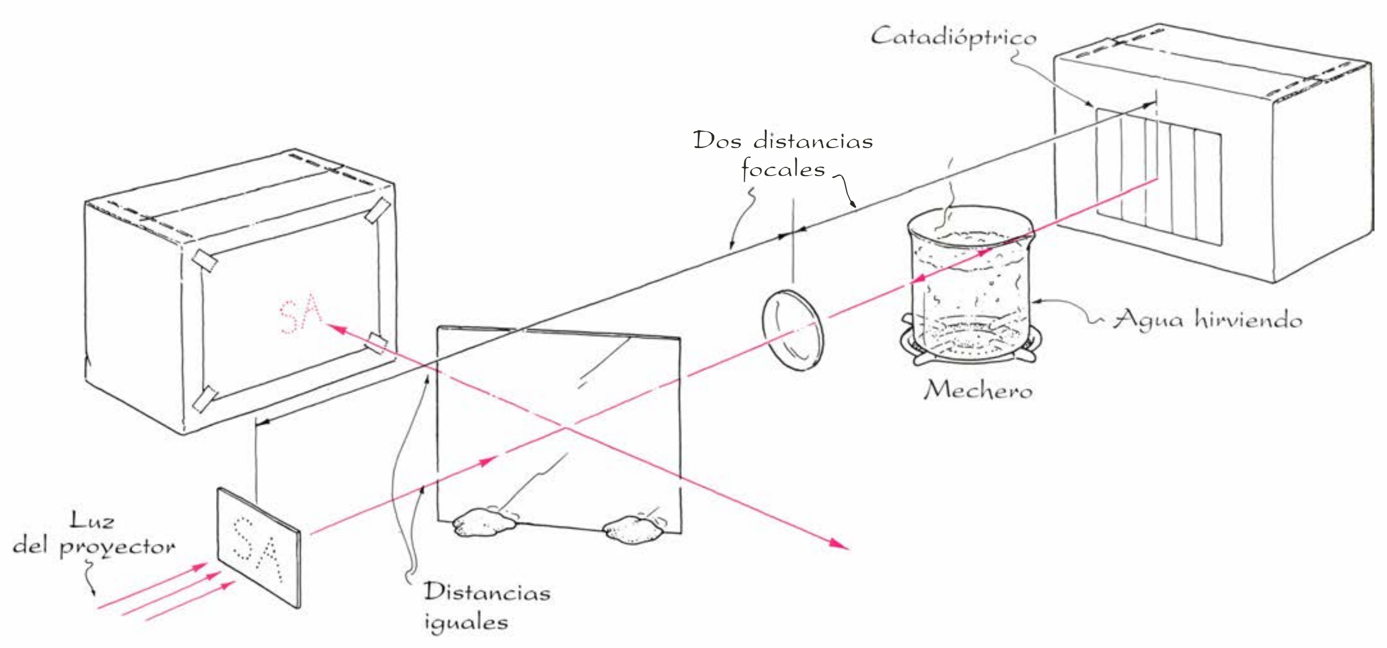
Punto inmóvil
Espejo plano unido al cono de un altavoz mediante un mango de madera
Lente a utilizar si se usa estenope
Luz del láser o del estenope del proyector
La vibración del espejo altera el trayecto de la luz
Separador de haces
Catadióptrico de cara al espejo plano
Señal de un oscilador de banda audio

7. Eliminación de la distorsión vibratoria en un haz

nocida en el espejo y la retro-reflexión, la cual hace que el haz vuelva sobre sí mismo.

Me entretuve en intercalar retículos de difracción y otros elementos perturbadores delante del catadióptrico. En todos los casos, con la retro-reflexión desaparece la distorsión sufrida por la luz en su primer paso por el elemento. Pero parte de la luz láser se refleja en la superficie delantera del elemento distorsionador y, por tanto, nunca lo atraviesa por segunda vez, con lo que esa distorsión no se elimina. Para obstaculizar la reflexión especular se gira el elemento distorsionador, junto con el catadióptrico, hasta que la reflexión aparezca en la cartulina; lo que quede entonces en la pantalla será la luz retroreflejada en el catadióptrico y que atravesó por segunda vez el elemento intercalado.

Con estos sistemas catadióptricos puede conseguirse la imagen de varios orificios. Para ello, se monta delante del proyector un ocultador en el que se haya perforado un conjunto de orificios dispuestos de modo que formen una o dos letras pequeñas. Con este dispositivo pueden repetirse todos los experi-



8. Experimento de David M. Pepper con la distorsión de turbulencia

mentos. Por ejemplo, cuando el elemento distorsionador y el catadióptrico están en posición, en la pantalla aparece una imagen borrosa de las letras. O sea, cada estenope crea su propia imagen y la combinación de todos ellos forma las letras.

Para mejorar la imagen se puede instalar una lente convexa entre el separador y el catadióptrico, mediada por dos distancias focales de los estenopes. Al otro lado de la lente, se coloca el catadióptrico a dos distancias focales. Según Pepper, la distancia focal no es crítica, pero el experimento sale mejor si su valor está entre 10 y 30 centímetros. Puede servir una simple lupa.

Para afinar la alineación se mueven la lente y el catadióptrico hasta que aparezca una imagen nítida e invertida en una hoja de papel colocada exactamente delante del catadióptrico. Se retira el papel y se mueve la pantalla hasta que la imagen resulte lo más nítida posible. Un objeto perturbador que se sitúe delante mismo del catadióptrico atenuará y emborronará la imagen, pero ésta seguirá viéndose. Para compensar la leve dispersión del haz que producen los retro-reflectores, en este experimento hace falta una lente. Este montaje podría permitir sustituir el ocultador de cartulina situado delante del proyector. Se retira la lente delantera y se introduce una diapositiva ordinaria en el proyector.

Pepper ha investigado asimismo de qué modo un catadióptrico puede neutralizar la distorsión producida por la vibración. Para ello, encoló un espejo plano de poco peso a un pequeño mango de madera, el cual unió a un cono de altavoz con cinta adhesiva engomada por las dos caras. Iluminó entonces el espejo con la luz que había atravesado el separador. (La fuente luminosa puede ser un láser, o bien un proyector provisto de ocultador con un estenope.) La luz reflejada en el espejo iba a parar a un catadióptrico. Seguidamente activó el altavoz con la señal amplificada procedente de un oscilador de audio. Cuando el espejo vibraba, el ángulo bajo el que el haz se reflejaba hacia el catadióptrico variaba continuamente. Pese a ello, la pantalla exhibía una imagen estacionaria de la apertura de salida del láser (o del estenope). El espejo, aun cuando vibraba con gran rapidez, se movía considerablemente más lento que la velocidad de la luz. Por eso, producía el efecto de permanecer inmóvil mientras la luz iba hasta el catadióptrico y venía. Al remplazar el catadióptrico por un espejo plano, Pepper observó en la pantalla una imagen que se movía en desorden.

Como última experiencia, Pepper eliminó la distorsión generada por una turbulencia. Para ello dispuso el montaje básico con un ocultador agujereado y un catadióptrico en un extremo del trayecto óptico. Asimismo, entre el separador y el espejo colocó una lente convexa, e instaló un mechero de gas apagado inmediatamente debajo del camino óptico, entre la lente y el espejo. (Podría haberlo puesto entre la lente y el separador.) El estenope, la pantalla y el catadióptrico estaban los tres a dos distancias focales de la lente. Tras afinar la alineación del sistema, encendió el mechero.

Cabría esperar que la imagen danzase y titilara al paso del haz por la turbulencia de la corriente convectiva creada por la llama. Y así sería con un espejo plano, pero en este caso permaneció quieta en su sitio. En efecto, aunque la deformación del haz variaba continuamente a causa de la turbulencia, el catadióptrico producía una imagen clara porque la luz pasaba dos veces por los puntos de perturbación antes que éstos tuvieran tiempo de cambiar.

Pepper colocó luego en el mechero un vaso de laboratorio con agua, interceptando el trayecto de la luz. Al calentarse el agua, la convección perturbaba el haz y, como antes, el catadióptrico eliminaba la distorsión. Cuando el agua comenzó a hervir, las burbujas impedían que parte de la luz llegase al catadióptrico y, en consecuencia, la imagen en la pantalla palideció algo.

Les brindo algo en qué pensar. Si los catadióptricos son espejos, ¿por qué no nos vemos retro-reflejados en uno de ellos cuando lo miramos bajo una luz ambiental normal? (Nos veremos si la habitación esta fuertemente iluminada y lo orientamos exactamente de frente, pero esa reflexión se debe a imperfecciones en la construcción.)

Probablemente alguien descubrirá nuevas posibilidades de los catadióptricos. Pepper y yo recibiremos con gusto sus noticias.

En febrero pasado les hablé de un calidoscopio equipado con filtros polarizados y una lámina reflectora curvada en su interior. Se trata del Karascope, encargado por el Museo de Arte Moderno de Nueva York, que inventó y patentó Judith Karelitz. El Museo lo vende al precio de 20,75 dólares, franco de portes. La dirección del Museo es 11 West Street, Nueva York, N.Y. 10019. No olviden consignar la correspondencia a la atención de Mail Order Department.

Libros

Origen de las razas humanas, cinética de reacciones, historia de la física, invierno nuclear y epistemología

T. Varela, J. Alonso, A. Roca, A. F. Rañada y M. Artigas

ADAPTACIONES RACIALES, por Carleton S. Coon. Editorial Labor; Barcelona, 1985. Cualquier tipo de investigación en el seno de las ciencias biológicas no presenta mayor número de dificultades que las inherentes a la propia disciplina en la cual queda enmarcado el estudio. Sin embargo, cuando los análisis se efectúan sobre la especie *Homo sapiens*, múltiples connotaciones incidirán en dichos estudios y determinarán agrias controversias, que dificultarán el mejor conocimiento de nuestra especie. Además, una serie de condicionamientos ético-sociales, exclusivos de la especie, han de ejercer una decisiva influencia en los planteamientos experimentales de las diferentes líneas de investigación.

La obra general de Carleton S. Coon entra en esa difícil maraña de intentar descifrar la realidad biológica de *Homo sapiens*. Como cualquier otra especie biológica, el hombre posee una variabilidad actual y una variabilidad en el pasado. Estas dos afirmaciones conllevan una serie de planteamientos previos; Coon considera, por una parte, que el hombre es el producto de un complejo proceso evolutivo, cuyo arranque se inicia en un conjunto de primates de características similares a los póngidos actuales: por otra, la variabilidad es el resultado de una serie de mecanismos genéticos en interacción con el medio. Para el autor, el medio es aquí cultura y clima. Ante los cambios ambientales, la especie responderá de una u otra forma, creándose unas variaciones que se distribuirán en el espacio. Esto es válido hasta un determinado momento de la historia evolutiva; posteriormente, la distribución de las características humanas en su relación con el espacio sufrirán importantes variaciones, que habrá de tener en cuenta la taxonomía. Basándose, esencialmente, en caracteres morfológicos establece cinco grandes razas o variedades dentro de *Homo sapiens*: Caucasoides, Mongoloides, Australoides, Negroides y Capoides, definida cada una de ellas por una serie de rasgos que la tipifican.

La taxonomía es la disciplina que se ocupa de la clasificación de los seres vivos. En la medida en que una entidad biológica, con variabilidad continua, se categoriza en un ente abstracto de límites precisos, interviene cierto grado de subjetivismo, campo abonado para la controversia. Así, han existido y existen bastantes discrepancias sobre la clasificación racial del *Homo sapiens*, hasta el punto de que algunos autores llegan incluso a negar la existencia de razas dentro de la especie humana. La tendencia dominante parece inclinarse hacia la aceptación de cuatro razas, agrupando en una sola entidad los Negroides y Capoides de Coon.

El autor se adentra en el análisis de la pigmentación del cabello, ojos, piel, etcétera. Nos ofrece un estudio exhaustivo de los procesos fisiológicos de la pigmentación y sus relaciones con múltiples constantes climáticas. En su opinión, las diferencias de pigmentación están muy relacionadas con las radiaciones solares, siendo las mismas una respuesta adaptativa de los individuos a climas diferentes. En la misma línea se plantea, y resuelve, el problema de la estatura.

Apartado crucial de la obra es el dedicado a *Homo erectus*. Coon pertenece a un grupo de antropólogos que defiende la teoría policentrista del origen de *Homo sapiens*. Para estos autores el estadio evolutivo del proceso de hominización, correspondiente al *Homo erectus*, presenta una importante variabilidad; ese polimorfismo lo dividen en cuatro o cinco variedades. Cada una de éstas evolucionaría independientemente y originarían las cuatro o cinco razas de *Homo sapiens*.

La teoría policentrista ha recibido duras críticas desde diversas áreas del campo de la biología. Una objeción muy sólida procede del campo de la genética de poblaciones. Según la misma, a partir de cuatro variedades de una especie, que evolucionan separadamente, la posibilidad de que las variedades se transformen en una misma especie, a su vez con cuatro variedades, presenta una probabilidad muy remota. El autor incorpora esas críticas a la teoría policentrista, pero no emite ningún juicio. Lo evidente en el proceso de hominización es la transformación del *erectus* en *sapiens*. ¿Cómo se produjo? La antropología física no tiene respuestas concluyentes.

Coon pasa revista pormenorizada a las formas fósiles posteriores al estadio de *Homo erectus* y anteriores al *Homo sapiens*. Plantea una serie de hipótesis que pretenden explicar algunos de los múltiples problemas tipológicos y filogenéticos derivados de la variabilidad de las formas y de la antigüedad de las mismas. Así, el relativo a la menor capacidad media del hombre actual con respecto a una etapa evolutiva anterior, el hombre de Neanderthal. En su afán globalizador aborda, entre otros, el oscuro origen de los primeros pobladores de América.

En la última parte del libro se detiene en la posible transmisión hereditaria de pautas de comportamiento en los mamíferos y en el hombre actual. Algunas de estas pautas de comportamiento arrancarían desde etapas tan ancestrales como la de los reptiles. Esta afirmación la funda en análisis de comportamiento comparado en diferentes grupos de animales, así como en el proceso evolutivo del hombre y en las razas actuales. (T. V.)

PROGRESS IN REACTION KINETICS. K. R. Jennings y R. B. Cundall (eds.). Pergamon Press; Oxford, 1983. Este libro constituye un nuevo volumen, el 11, de una prestigiosa serie dedicada a artículos de puesta a punto sobre un importante área de la química-física: la cinética de reacciones. Consta de tres artículos sobre tres diferentes temas.

El primero de ellos, "Procesos de colisión en el láser de fluoruro de hidrógeno", está escrito por R. D. H. Brown y A. Maitland, de la Universidad de St. Andrews en Escocia. Los estudios sobre el láser de fluoruro de hidrógeno (HF) abordan normalmente el funcionamiento práctico de prototipos y los modelos de láser obtenidos por simulación por computador. Algunos investigadores han considerado la relación entre ambos estudios y han comprobado que el acuerdo es sólo imperfecto.

Esta situación levantó la sospecha de que algunos valores hasta entonces aceptados de las constantes de velocidad para ciertas reacciones químicas implicadas en el funcionamiento de dicho láser pudieran no ser correctas, lo que ha originado una proliferación de trabajos sobre medidas experimentales de constantes de velocidad. A la vez, los modelos teóricos han ido refinándose. El objetivo primordial del trabajo de Brown y Maitland es proporcionar una compilación de constantes de velocidad experimentales para procesos de colisión importantes en el láser de HF. En particular, se ocupan de la cinética de formación y eliminación de posibles moléculas que den lugar al efecto láser. La determinación de constantes de velocidad cubre una importante área experimental, no sólo por su interés para los modelos de láser sino también por su uso en el área de la dinámica molecular. Una parte del trabajo se refiere a dos reacciones de bombeo: $F + H_2 \leftrightharpoons HF\ (r,v) + H$ y $H + F_2 \leftrightharpoons HF\ (r,v) + F$; en éstas se crean moléculas HF en estados rotacionales (r) y vibracionales (v) excitados, es decir, una inversión de poblaciones.

Los autores presentan no sólo los resultados para las constantes de velocidad sino que describen también y examinan las técnicas experimentales usadas para medirlas. De hecho, la primera de las reacciones arriba escritas es una de las mejor estudiadas en química. La segunda parte del trabajo se refiere a reacciones que dan lugar a la relajación rotacional y vibracional de las poblaciones obtenidas mediante las reacciones de bombeo. Mediante choques de las moléculas HF con átomos o moléculas, la población inicial cambia. Esta alteración tiene gran importancia para determinar qué transiciones originarán el efecto láser y la energía producida. Se dedica en el trabajo un gran esfuerzo a aquellos procesos de relajación vibracional que puedan limitar el rendimiento del láser.

El segundo artículo se titula "El papel de los radicales libres en la carcinogénesis química y por radiación" y está escrito por C. L. Greenstock, del Whiteshell Nuclear Research Establishment en Manitoba, Canadá. Si este es un trabajo fascinante para el no experto en bioquímica, ¡cuánto más lo será para el experto! El artículo está muy bien estructurado y constituye una interesante contribución a la investigación de las bases moleculares del cáncer. Una excelente introducción sirve para excitar la curiosidad del lector. No me resisto a exponer brevemente las principales ideas. Hoy día, alrededor del veinte por ciento de la humanidad muere de cáncer. Una gran parte se debe a efectos naturales del medio ambiente: radiación solar, radiación natural de fondo (terrestre o procedente de fuentes cósmicas) y otra a causas producidas por el hombre, como la polución originada por industrias, automóviles o viviendas, la acumulación de desperdicios tóxicos, aditivos en los alimentos, etcétera. La radiación solar es la responsable del cáncer de

piel. Entre los debidos a la intervención humana destaca el cáncer de pulmón, cuya incidencia se correlaciona fuertemente con el hábito de fumar. La edad incide también, por la menor resistencia celular a los ataques de substancias carcinógenas con los años.

Los radicales libres son intermediarios responsables de iniciar la mayoría de los efectos biológicos nocivos de los carcinógenos físicos (radiación) o químicos. Estos radicales libres son fragmentos moleculares que contienen un electrón no apareado en un orbital electrónico particular, electrón solitario que busca otro electrón con el cual aparearse y adquirir una configuración menos reactiva. El proceso de robar un electrón de una molécula vecina estable y oxidarla es la manera en que los radicales libres manifiestan sus efectos biológicos perniciosos. Hay bastantes indicios de la implicación de formas diversas de oxígeno activado en dichos efectos biológicos nocivos mediante promoción o catálisis de los procesos de oxidación por radicales libres. El trabajo trata con detalle y claridad la acción de los radicales libres, así como los procesos de activación, ya que muchas sustancias potencialmente carcinógenas sólo lo son después de ser activadas metabólicamente por procesos bioquímicos o por radiación solar.

No se olvida de la protección contra el cáncer. La primera medida preventiva consiste en disminuir la exposición a potenciales agentes carcinógenos. La segunda línea de defensa se desarrolla en el nivel celular. Consiste en reducir la cantidad de carcinógenos activados estimulando la propia respuesta celular para autodesintoxicarse (una alimentación adecuada contribuye a ello) o bien por medicación externa (quimioterapia y radioterapia). La investigación farmacológica sobre drogas antitumorales que desactiven los radicales libres está descubriendo algunos de los métodos de autoprotección de las propias células.

El tercer artículo lleva por título "Teoría de la interacción entre iones y especies neutras. Aplicación de la teoría del estado de transición a propiedades de colisión y reactivas de sistemas simples"; está escrito por W. J. Chesnavit y M. T. Bowers, de la Universidad de California en Santa Bárbara. El trabajo, más extenso que los anteriores, entra a fondo en el tejido matemático. El área de la química de iones en fase gaseosa ha experimentado un crecimiento espectacular durante los últimos veinte años. Este crecimiento ha sido posible gracias a las innovaciones técnicas en electrónica y en la tecnología del vacío, que ha permitido el desarrollo de una instrumentación refinada. Aunque la mayor parte del progreso ha sido experimental, también se han dado importantes avances en la teoría cinética. Estos avances teóricos se refieren sobre todo a sistemas neutros y no a sistemas iónicos.

Gran parte de la teoría básica desarrollada para sistemas neutros puede adaptarse a iones, pero sólo después de sustanciales modificaciones. Este es el tema del trabajo de Chesnavit y Bowers. El artículo se divide en dos partes. La primera expone la teoría de colisiones. La teoría del estado de transición proporciona el marco para el trabajo. La segunda parte trata de los sistemas reactivos. Se explora la naturaleza del estado de transición y se presentan modelos que cubren una amplia gama de sistemas reactivos. En ambas secciones la teoría se presenta de tal forma que tanto el experto como el no experto pueden entender las hipótesis y las aplicaciones de los conceptos a sistemas reales (en el supuesto de que el lector posea un conocimiento adecuado de mecánica cuántica). Las secciones sobre teoría de reacciones demuestran la necesaria simbiosis entre teoría y experimento.

En resumen, nos hallamos ante un libro que debería estar en las bibliotecas de nuestras facultades de química o física, así como el resto de los volúmenes de la serie "Progress in Reaction Kinetics", a la que el presente volumen pertence. (J. A.)

HISTORIA DE LA FÍSICA HASTA EL SIGLO XIX. Real Academia de Ciencias Exactas, Físicas y Naturales; Madrid, 1983. El volumen recoge los textos de las conferencias de un curso con el mismo título organizado por la institución científica madrileña durante los meses de abril y mayo de 1983. Sus autores son Juan Vernet, Julio Samsó, Eduardo Millás Vendrell, Carlos Solís, Alberto Elena, Manuel García Doncel, Armando Durán, Antonio Moreno, Alberto Dou, José M. Torroja, Máximo Rodríguez Vidal, David Jou y Víctor Navarro. Por este orden, los conferenciantes tratan aspectos de la historia de la física desde la antigüedad hasta el siglo XVIII. Carlos Sánchez del Río, organizador de las conferencias, advierte que se dejan los siglos XIX y XX para otros cursos.

La edición de este volumen tiene una importancia obvia. La bibliografía española relativa a la historia de la ciencia, pobre aún, no conocía ninguna obra similar; hay que mencionar, sin embargo, traducciones de obras de contenido paralelo, la mayoría nacidas por la iniciativa del grupo hispanoamericano (argentino, sobre todo) de historia de la ciencia en las décadas de los cuarenta y los cincuenta. La historia de la física, a la que algunos de los autores del presente volumen han contribuido con trabajos originales y de investigación, debe buscarse, en España, dispersa en revistas y volúmenes colectivos varios.

El objetivo de la mayoría de los conferenciantes ha sido, según puede verse leyendo los distintos textos, hacer llegar a nuestro país los resultados de una disciplina, la historia de la física, que ha conocido en los últimos años un desarrollo notable. El tema del curso, la física "internacional" (excepto en las dos conferencias de Víctor Navarro), lleva a un tratamiento bibliográfico de valor, en su conjunto, y, sólo en algunas ocasiones, se refiere a investigaciones originales. Quizá, no podía ser de otro modo. Nuestros especialistas, en algún caso de renombre mundial, deben compaginar sus investigaciones con ocupaciones profesionales no siempre acordes con su actividad científica. El gesto –quizá simbólico– de la Real Academia de Ciencias, organizando cursos de historia de la ciencia como el que merece el presente comentario, deberían tenerlo en cuenta otros sectores del mundo científico español, donde la historia de la ciencia sigue siendo una actividad propia de jubilados (algunos de los cuales, hay que decirlo, han hecho interesantes contribuciones) o adecuada para personas no aptas para la "verdadera" investigación, que se define, paradójicamente, por la periclitada división de los saberes que refleja aún nuestra universidad. Hay que decir también que algunas instituciones parecen estar cambiando de actitud; mencionemos, por ejemplo, el Institut d'Estudis Catalans y algunos centros superiores.

La física fue, después de la astronomía, la "ciencia filosófica" por excelencia y provocó reflexiones de distinto carácter desde su formulación independiente, sobre todo a cargo de Isaac Newton. No es de extrañar, pues, que los estudios históricos de la ciencia se hayan ocupado de la física (y de la astronomía) de manera muy destacada, aunque los trabajos sobre historia de las matemáticas sean también de valor

(muchas veces, sin embargo, ligados a la historia de la física y de la astronomía). En el panorama de los estudios actuales puede verse, sin embargo, una gran diversificación de enfoques. Aunque los estudios ya clásicos de historia de la física, por ejemplo los de Pierre Duhem, trataban de forma casi exclusiva la evolución de las *ideas* físicas, hoy empiezan a aparecer otras corrientes tendentes a situarlas en su contexto social (económico, político y cultural).

El volumen de la Academia recoge ambas líneas fundamentales de la historia de la física y en la proporción en que se encuentran en la literatura universal; predominan lógicamente los estudios "internos" de la física (era imprescindible, por ejemplo, haber hecho una edición crítica de los *Principia* de Newton). Las dos intervenciones de Víctor Navarro sobre el cultivo de la física en España en los siglos XVI-XVIII responden a la investigación más "acabada" de la difusión de las ideas físicas en la historia de España que, hasta el momento, podemos disponer. Sin que se haya agotado este tema, otros períodos de la historia permanecen casi vírgenes. A pesar de ello, un lector atento observará, en varias de las intervenciones, referencias a la historia de la física en España, aunque lo tratado no sea específicamente esta cuestión. La obra aborda los puntos siguientes: la física en la antigüedad, en el mundo arábigo-islámico, en el mundo latino medieval, la mecánica de Galileo, los precursores de Newton, un extenso comentario de los *Principia* de Newton, la óptica y la física de fluidos en el siglo XVII, la mecánica, la mecánica celeste, la electricidad y la termología en el siglo XVIII. Se trata, pues, de aproximaciones de conjunto en períodos o en disciplinas concretas. Los autores, no todos ellos investigadores "profesionales" en historia de la ciencia (aunque, por los comentarios anteriores, esta demarcación sea poco rigurosa hoy día entre nosotros), se benefician, en general, de las aportaciones más significativas de la disciplina en otras regiones del mundo. (A. R.)

NUCLEAR WINTER, por O. Greene, I. Percival e I. Ridge. Polity Press, 1985. En 1955, A. Einstein y B. Russell afirmaron, en su famoso manifiesto sobre los peligros de una guerra nuclear, que "tenemos que aprender a pensar de un modo nuevo" si queremos evitar que una confrontación entre los dos bloques acabe con la vida humana sobre la tierra.

Desde entonces son muchos los científicos que han hablado claramente y que se han opuesto a ella, junto con grandes sectores de opinión. Pero, a pesar de ello, los arsenales nucleares han evolucionado de manera constante. Pues no sólo ha crecido extraordinariamente el número y la potencia de las cabezas nucleares (que equivale hoy a más de un millón de bombas de Hiroshima), sino que éstas son cada vez más precisas, operativas e invulnerables. Desde la fabricación de la bomba H, en los años 50, la carrera de armamentos está jalonada por muchos saltos cualitativos importantes, como la aparición de los proyectiles balísticos intercontinentales, de los lanzados desde submarinos, de los vehículos de reentrada múltiple o de muchos tipos de armas tácticas; hoy mismo asistimos a uno más con el desarrollo de las armas antisatélites y de la ISD, popularmente conocida como guerra de las galaxias.

Pensar de un modo nuevo es aún más necesario hoy que en 1955. Pero muchos políticos y líderes de opinión, así

como sectores mal informados de sus electores, siguen pensando que una guerra nuclear sólo se diferencia de las anteriores por su intensidad. Creen que habría más muertos y más destrucción, pero que la humanidad podría recuperarse, como lo hizo en el caso de las grandes guerras del pasado. Incluso han indicios que permiten asegurar que se especula sobre el número de muertos, por debajo del cual una guerra nuclear podría ser rentable para alguna de las superpotencias. Sin duda no han comprendido que una guerra nuclear sería una confrontación de un tipo radicalmente nuevo, con un nivel de destrucción, muerte y sufrimientos tan alto, que toda referencia a los conflictos anteriores no tiene ningún valor.

El manifiesto Einstein-Russell puso en acción a numerosos científicos de todo el mundo para estudiar, de manera serena e independiente, las consecuencias del uso de armas nucleares. Surgió así el movimiento Pugwash, al que pronto seguirían otras asociaciones de personas cuya cualificación les permite abordar los aspectos técnicos de la cuestión. Sin duda, en la toma de decisiones todos los votos deben valer lo mismo, según el claro principio democrático. Pero es necesario que, en la información previa a la opinión pública, los científicos desempeñen un papel especial para que las decisiones se basen en un conocimiento correcto.

Todos esos esfuerzos fueron conduciendo a una consecuencia clara: con una probabilidad muy alta, una guerra nuclear, de cualquier tipo que fuese, se escaparía de todo control y produciría un cataclismo de tal magnitud que podría terminar con la especie humana. Se justifican así las dramáticas palabras de Einstein y Russell, quienes hablaban "no como miembros de esta o aquella nación, continente o credo, sino como seres humanos, miembros de la especie hombre, cuya existencia continuada está en duda" y quienes terminaban su llamamiento apelando "como seres humanos a seres humanos: recordad la humanidad y olvidad el resto.. si no podéis hacerlo, el riesgo de la muerte universal está ante vosotros".

Esta terrible advertencia ha recibido un respaldo muy fuerte en los últimos dos años, con el descubrimiento de que a los ya conocidos tres efectos de una guerra nuclear (las ondas explosivas, las bolas de fuego y la radiactividad) hay que añadir una probable perturbación catastrófica del clima terrestre. Se trata del llamado invierno nuclear, período frío y oscuro, con violentos vientos y nieblas tóxicas, que alargaría los efectos del primer momento al interrumpir la producción de alimentos y dificultar grandemente la reconstrucción de la economía y de la estructura social, aumentando de tal modo los sufrimientos de los supervivientes que podría llevar por sí mismo a una lenta extinción de la vida humana.

Se estima generalmente que los efectos inmediatos de una guerra nuclear a gran escala producirían la muerte de entre el 70 y el 80 por ciento de la población europea en un plazo de unos pocos días. Se sabía que los supervivientes estarían sometidos a una intensa y dañina radiación residual de largo plazo. Pero hasta 1982 no se empezó a comprender que una amenaza, que podría ser mayor, les aguardaría. En ese año, el holandés P. Crutzen y el norteamericano J. Birks realizaron, por encargo de la revista sueca de ciencias ambientales *Ambio*, un trabajo sobre los efectos de una guerra nuclear sobre la atmósfera. Aunque la idea original de *Ambio* se refería al aumento de radiación ultravioleta, ellos pensaron que sería interesante estudiar el efecto de los humos producidos. Tras unas estimaciones simples, encontraron asombrados que el humo producido bloquearía casi toda la luz del sol durante varias semanas al menos.

A raíz de la publicación de sus conclusiones y resultados, cinco norteamericanos, especialistas en física atmosférica y planetaria, R. P. Turco, O. B. Toon, T. P. Ackerman, J. B. Pollack y C. Sagan, desarrollaron un modelo más complejo y detallado y encontraron que el verano podría transformarse en invierno, confirmando así las conclusiones de Crutzen y Birks (véase un informe de esos cinco autores en *Investigación y Ciencia*, octubre de 1984). Impresionados por su predicción, hicieron un llamamiento a los científicos de todo el mundo para estudiar el problema y examinar sus resultados. Un grupo de 20 biólogos, dirigidos por P. Ehrlich, de Stanford, aceptaron la invitación, llegando rápidamente a concluir en 1983 que muchas especies animales y vegetales podrían extinguirse, de manera parecida a como ocurrió con los dinosaurios y otros seres, cuya desaparición fue, sin duda, debida a acontecimientos catastróficos que modificaron el clima de la tierra hace millones de años.

Numerosos equipos científicos e instituciones de todo el mundo empezaron a comunicar resultados de estudios sobre el nuevo tema. Entre ellos destaca el detallado y exhaustivo análisis del National Research Council de la Academia de Ciencias de los Estados Unidos, dado a conocer en 1985. Los datos científicos acumulados en estos tres años son tajantes: una guerra nuclear a gran escala produciría una alteración brusca del clima terrestre, sin comparación posible con ninguno de los fenómenos ocurridos desde la aparición de la vida humana en la tierra. Un período frío y oscuro de al menos varios meses podría llegar a impedir casi totalmente la producción de alimentos en el hemisferio norte y a disminuirla gravemente durante varios más.

Greene, Percival y Ridge, que trabajan desde hace varios años en organizaciones de científicos contra la guerra nuclear, han escrito un libro muy interesante y oportuno. Aunque se trata de un tema difícil y complejo, han ordenado eficazmente una gran cantidad de información. En el cuerpo principal del libro, formado por nueve capítulos y 170 páginas, exponen de manera clara la situación de los arsenales nucleares, las propiedades dinámicas de la atmósfera relevantes para el problema, la evolución previsible de los humos y los efectos del invierno nuclear sobre los seres vivos. En 10 apéndices reúnen algunos cálculos que, sin ser muy complejos, tienen algún detalle técnico. Finalmente, una cuidadosa bibliografía remite al lector con suficiente formación científica a las obras originales. Esta estructura obedece a la finalidad básica de la obra, que es la información de la opinión pública sobre las consecuencias de una guerra nuclear. Ciertamente consigue su objetivo, pues se puede leer en varios niveles: por personas sin formación científica especializada para comprender el estado de la cuestión o como iniciación a su estudio riguroso utilizando la bibliografía recogida.

La intensidad del invierno nuclear depende mucho del momento del año en que se produce la guerra, siendo mayor si se inicia en verano y menor en invierno. En el primer caso, sobre todo si tiene lugar en junio, serían mucho mayores tanto los incendios como los efectos sobre las cosechas. El libro se concentra en una guerra de verano, siguiendo lo que se considera normalmente como uno de los guiones más probables. En él, una agravación de las tensiones en Oriente Medio provoca la puesta en estado de alerta roja de las fuerzas de la OTAN y del Pacto de Varsovia. Como precaución ante un ataque preventivo, cada parte dispersa

sus cabezas nucleares con base en submarinos y otros buques, así como los misiles crucero. Estas maniobras son detectadas e interpretadas como preparativos de un ataque nuclear y, para evitarlo, uno de los bloques inicia uno con armas convencionales. De manera muy rápida, se produce una escalada que lleva a la utilización de armas nucleares tácticas en Europa. Pronto son afectados los distritos más occidentales de la URSS; además, los jefes militares de Francia y el Reino Unido, cuya fuerza nuclear quedaría muy disminuida en los primeros momentos, utilizarían sus últimas cabezas contra blancos en Rusia. Las dos superpotencias aceptarían finalmente que una guerra a gran escala es inevitable y lanzarían sus proyectiles balísticos intercontinentales contra sus territorios y sus bases e instalaciones militares en todo el mundo.

En esa hipótesis se utilizarían unas 18.000 cabezas nucleares, con una potencia de unos 6000 megatones, equivalente a cerca de medio millón de bombas de Hiroshima, pero sólo un 40 por ciento del arsenal existente. La mayoría de las explosiones se producirían entre el círculo polar ártico y el trópico de Cáncer, donde vive la mitad de la población mundial.

El invierno nuclear se produciría como efecto de los humos de los incendios. Según los autores, la extensión de la zona incendiada en Europa sería superior al millón de kilómetros cuadrados, con tres tipos distintos de fuegos. Las zonas urbanas quemadas tendrían una superficie similar a la de la unión de Italia y Suiza, las de bosques más de cinco veces la de Benelux y las de cultivos dos veces la de Gran Bretaña.

En varios capítulos se expone cuáles son los mecanismos básicos de circulación del aire que permiten comprender lo que ocurriría con las partículas que componen el humo. Parte caería al suelo al cabo de horas o pocos días. Pero las que superasen la tropopausa y entrasen en la estratosfera, caracterizada por haber en ella muy pocos movimientos verticales, quedarían allí durante mucho tiempo, distribuyéndose de manera aproximadamente homogénea sobre toda la banda de superficie de la misma latitud. Se formaría así una película que cortaría el paso a los rayos solares, primero en las zonas templada y tropical del hemisferio norte, y luego probablemente se extendería al hemisferio sur. El cálculo de este proceso es complicado y difícil, por lo que los autores dedican un capítulo a examinar críticamente las incertidumbres y posibles puntos débiles de los razonamientos.

También exponen de manera clara y resumida el efecto que los cambios de clima tienen sobre los animales y plantas, con la consecuencia de que la producción de alimentos podría anularse completamente durante 10 meses en la zona templada septentrional, tardándose varios años en la recuperación de las cosechas, si ésta llega a producirse. Aunque el libro se concentra en una guerra como la descrita más arriba, los autores estudian también otras varias posibilidades: una guerra más suave, otra más fuerte y una de invierno. Finalmente, en su último capítulo exponen las consecuencias que la comprobación de este peligro debería tener sobre las decisiones políticas. (A. F. R.)

¿Salvar lo real? Materiales para una Filosofía de la Ciencia, por Alfonso Pérez de Laborda. Ediciones Encuentro; Madrid, 1984. El autor, profesor de historia de la ciencia en Salamanca, había ya sobresalido por un estudio monográfico sobre Leibniz y Newton y otro libro sobre aspectos ideológicos de la ciencia. Recoge aquí 12 trabajos sobre puntos diversos de la filosofía de la ciencia: verdad, probabilidad, inducción, materialismo, ciencia y creencia. Pérez de Laborda defiende el carácter realista del conocimiento científico y critica los ingenuos reduccionismos. "El pensamiento va construyendo la realidad, pero en la misma construcción –porque *aquí* la aproximación a la realidad se hace por el pensamiento y con él, y no de alguna otra manera distinta– nos aparece la inmensa realidad como inabarcable y como configurativa, como inasible y como misteriosa, que escapa siempre más y más a nuestras estrategias conocedoras. Y esto lo conocemos: he aquí el punto clave que quiere expresarse a lo largo y a lo ancho de este libro."

Fino analista, expone con claridad y en un lenguaje llano planteamientos de la epistemología moderna: neopositivismo, Popper, Stegmüller, la escuela de Erlangen, Hübner y demás con su evolución teórica, así como los ecos en algunos autores hispanos. Su construcción es segura, firmemente trabada, propia del ingeniero que es por formación antes de llegar a la filosofía. El lector me va a permitir la recomendación de las páginas dedicadas a las interpretaciones frecuencialista y subjetiva de la probabilidad. (M. A.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

COMUNICACIONES OPTICAS POR CABLES SUBMARINOS

BELL SYSTEMS TECHNICAL STAFF en *Bell System Technical Journal* (número monográfico dedicado a cables submarinos), vol. 57, n.º 7; 1978.

IEEE JOURNAL ON SELECTED AREAS IN COMUNICATION, vol. SAC-2, n.º 6 (número monográfico dedicado a cables submarinos por fibra óptica); noviembre, 1984.

COMUNICACIONES POR FIBRA OPTICA: MANUAL DE INGENIERÍA. R. Díaz de la Iglesia. Marcombo; Barcelona, 1985.

DINAMICA DE LAS PROTEINAS

DYNAMICS OF FOLDED PROTEINS. J. A. McCammon, B. R. Gelin y M. Karplus en *Nature*, vol. 267, n.º 5612, págs. 585-589; 16 de junio de 1977.

DYNAMICS OF PROTEINS. P. G. Debrunner y H. Frauenfelder en *Annual Review of Physical Chemistry*, vol. 33, págs. 283-299; 1982.

DYNAMICS OF PROTEINS: ELEMENTS AND FUNCTION. M. Karplus y J. A. McCammon en *An. Rev. of Bioch.*, vol. 52, págs. 263-300; 1983.

MAPPING PROTEIN DYNAMICS BY X-RAY DIFFRACTION. Dagmar Ringe y Gregory A. Petsko en *Progress in Biophysics and Molecular Biology*, vol. 45, n.º 3, págs. 197-235; 1985.

APLICACIONES DEL EFECTO HALL CUANTICO

THEORY OF THE QUANTIZED HALL CONDUCTANCE. B. I. Halperin en *Helvetica Physica Acta*, vol. 56, n.º 1-3; págs. 75-102; 1983.

THE QUANTIZED HALL EFFECT. H. L. Stormer y D. C. Tsui en *Science*, n.º 4603, págs. 1241-1246; 1983.

PROCEEDINGS OF THE FIFTH INTERNATIONAL CONFERENCE ON ELECTRICAL PROPERTIES OF TWO-DIMENSIONAL SYSTEMS, HELD AT OXFORD UNIVERSITY, 5-9 SEPTEMBER, 1983. Dirigido por R. J. Nicholas en *Surface Science*, vol. 142; 1984.

VON KLITZING WINS NOBEL PRIZE FOR QUANTUM HALL EFFECT. B. Schwarzschild en *Physics Today*, vol. 38, n.º 12, págs. 17-20; diciembre, 1985.

RITMOS INTERNOS EN LA EMIGRACION DE LAS AVES

CIRCANNUAL SYSTEMS. E. Gwinner en *Handbook of Behavioral Neurobiology*, vol. 4, dirigido por J. Aschoff. Plenum Press, 1981.

BIRD MIGRATION. Chris Mead. Facts on File Publications, 1983.

CIRCANNUAL RHYTHMS IN THE CONTROL OF AVIAN MIGRATIONS. E. Gwinner en *Advances in the Study of Behavior*, vol. 16, págs. 191-228; 1986.

NUEVAS ESTRATEGIAS PARA LA ENERGIA NUCLEAR

NUCLEAR POWER IN AN AGE OF UNCERTAINTY. Oficina de Valoración Tecnológica del Congreso de los Estados Unidos. Servicio de Publicaciones del Gobierno de los Estados Unidos, 1984.

NATIONAL STRATEGIES FOR NUCLEAR POWER REACTOR DEVELOPMENT. Richard K. Lester, Michael J. Driscoll, Michael W. Golay, David D. Lanning y Lawrence M. Lidsky. Departamento de Ingeniería Nuclear del Instituto de Tecnología de Massachusetts; marzo, 1985.

GIROSCOPIOS OPTICOS

MULTIOSCILLATOR LASER GYROS. W. W. Chow, J. B. Hambenne, T. J. Hutchings, V. E. Sanders, M. Sargent III y M. O. Scully en *IEEE Journal of Quantum Electronics*, vol. QE-16, n.º 9, págs. 918-936; septiembre, 1980.

THE RING LASER GYRO. W. W. Chow, J. Gea-Banacloche, L. M. Pedrotti, V. E. Sanders, W. Scleich y M. O. Scully en *Reviews of Modern Physics*, vol. 57, n.º 1, págs. 61-104; enero, 1985.

GYROSCOPES MAY CEASE SPINNING. Graham J. Martin en *IEEE Spectrum*, vol. 23, n.º 2, págs. 48-53; febrero, 1986.

LAS TENSIONES DEL NACIMIENTO

CATECHOLAMINE RELEASE IN THE NEWBORN. H. Lagercrantz y P. Bistoletti en *Pediatrics Research*, vol. 11, n.º 8, págs. 889-893; agosto, 1977.

OF LABOUR AND THE LUNGS. R. E. Olver en *Archives of Disease in Childhood*, vol. 56, n.º 9, págs. 659-662; septiembre, 1981.

FETAL SCALP CATECHOLAMINES DURING LABOR. Peter Bistoletti, Lars Nylund, Hugo Lagercrantz, Paul Hjemdahl y Helge Ström en *American Journal of Obstetrics and Gynecology*, vol. 147, n.º 7, págs. 785-788; 1 de diciembre de 1983.

SYMPATHOADRENAL ACTIVITY AND PERIPHERAL BLOOD FLOW AFTER BIRTH: COMPARISON IN INFANTS DELIVERED VAGINALLY AND BY CESAREAN SECTION. G. Faxelius, H. Lagercrantz y A. Yao en *Journal of Pediatrics*, volumen 105, n.º 1, págs. 144-148; julio, 1984.

INTERPRETACION DE LA TROMPETA BARROCA

THE MUSIC AND HISTORY OF THE BAROQUE TRUMPET BEFORE 1721. Don L. Smithers. Syracuse University Press, 1973.

ACOUSTICS AND THE PERFORMANCE OF MUSIC. Jürgen Meyer. Verlag Das Musikinstrument, Frankfurt, 1978.

REGENERATION IN BRASS WIND INSTRUMENTS. S. J. Elliott y J. M. Bowsher en *Journal of Sound and Vibration*, vol. 83, n.º 2, págs. 181-217; 22 de julio de 1982.

JUEGOS DE ORDENADOR

THE FOURTH DIMENSION SIMPLY EXPLAINED: A COLLECTION OF ESSAYS SELECTED FROM THOSE SUBMITTED IN THE SCIENTIFIC AMERICAN'S PRIZE COMPETITION. Dirigido por Henry P. Manning. Dover Publications, Inc., 1960.

PLANILANDIA. Edwin A. Abbott. Edit. Labor; Barcelona, 1976.

THE FOURTH DIMENSION: TOWARD A GEOMETRY OF HIGHER REALITY. Rudy Rucker. Houghton Mifflin Company, 1984.

THE SHAPE OF SPACE: HOW TO VISUALIZE SURFACES AND THREE-DIMENSIONAL MANIFOLDS. Jeffrey R. Weeks. Marcel Dekker, Inc., 1985.

TALLER Y LABORATORIO

RETROREFLECTIVE ARRAYS AS APPROXIMATE PHASE CONJUGATORS. H. H. Barrett y S. F. Jacobs en *Optics Letters*, volumen 4, n.º 6, páginas 190-192; 1979.

EXPERIMENTS WITH RETRODIRECTIVE ARRAYS. Stephen F. Jacobs en *Optical Engineering*, vol. 21, n.º 2, págs. 281-283; marzo/abril, 1982.